DOCUMENT RESUME

ED 060 895 LI 003 563

AUTHOR	Mackenzie, A. Graham, Ed.; Stuart, Ian M., Ed.
TITLE	Planning Library Services; Proceedings of a Research Seminar Held at the University of Lancaster 9-11 July 1969.
INSTITUTION	Lancaster Univ. (England). Library.
PUB DATE	69
NOTE	251p.;(117 References); Occasional Paper 3
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	Foreign Countries; *Libraries; Library Research; *Library Services; *Mathematical Models; *Planning; Research; Seminars; Techniques; *University Libraries
IDENTIFIERS	England; *University of Lancaster

ABSTRACT

This proceedings volume of a seminar on planning library services is the third in a series of papers, published at irregular intervals, to report on research work by members of the University of Lancaster library staff. From January 1967 until June 1969 the Office of Scientific and Technical Information (OSTI) organized regular meetings under the title of Operational Studies Seminars to identify and describe in quantitative terms various factors which affect the operations of libraries especially university libraries. There were no formal records of these meetings, and after OSTI terminated its support, the University of Lancaster research team organized its own meetings and published the proceedings. The aim is to see how far it is possible to carry out mathematical modelling of the operation of an academic library and its interactions with users, so that managerial decisions can be made on a rational rather than an institutional basis. The broad subject areas covered are: (1) Problems of library planning; (2) Techniques of analysis; (3) Data collection and evaluation; and (4) A review covering Library Management Research and Planning Library Services: an overview. This volume contains preprints of the papers presented and edited versions of the discussions. A list of participants is also included. This document previously announced as ED 045 173. (SG)

ED 060895

University of Lancaster Library Occasional Papers, No. 3

PLANNING LIBRARY SERVICES

Proceedings of a Research Seminar held at the University of Lancaster 9 – 11 July 1969

Edited by A. Graham Mackenzie
and Ian M. Stuart

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY University of Lancaster TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

LI 003 563

University of Lancaster Library

1969

238
117

PREFACE

The Proceedings of this Seminar form the third of a series of papers which we are publishing at irregular intervals to report on research work by members of the Library staff (see inside back cover).

From January 1967 until June 1969 a research programme was supported by the Office for Scientific and Technical Information; it has now become part of the Library's normal activities. The aim of this work is to see how far it is possible to carry out mathematical modelling of the operation of an academic library and its interactions with users, in order that managerial decisions may be made on a rational rather than on an intuitional basis. The research team comprised Mr. Michael K. Buckland, Dr. Anthony Hindle and Mr. Ian Woodburn.

CONTENTS

PREFACE

CONTENTS

FOREWORD

SEMINAR MEMBERSHIP

SESSION 1: PROBLEMS OF LIBRARY PLANNING

H. F. Dammers	The library/information service in an industrial research establishment: planning and prospects
F. F. Leimkuhler	On information storage models
A. Longworth	Public library services
D. J. Urquhart	National libraries
Discussion	

SESSION 2: TECHNIQUES OF ANALYSIS

C. F. Carter	The allocation of resources in higher education
R. M. Morley	Maximising the benefit from library resources
Discussion	

SESSION 3: TECHNIQUES (Continued)

B. C. Brookes	Statistical distributions in documentation and library planning
F. F. Leimkuhler	Storage policies for information systems
I. Woodburn	A mathematical model of a hierarchical library system
Discussion	

SESSION 4: TECHNIQUES (Continued)

M. K. Buckland & A. Hindle	Loan policies, duplication and availability
Discussion	

SESSION 5: DATA COLLECTION AND EVALUATION

R. M. Duchesne	Library management information from computer-aided library systems
M. G. Ford	Data collection and feedback
A. Hindle	Models and measures for non-profit making services
Discussion	

SESSION 6: REVIEW

B. C. Vickery	Library management research: what next?
F. F. Leimkuhler	Planning library services: an overview
Discussion	

FOREWORD

For the past two or three years work has been proceeding within the Universities of Durham and Lancaster and elsewhere to identify and describe in quantitative terms various factors which affect the operation of libraries of all types, but principally within the university sphere. OSTI, which has provided financial support for many of these programmes, has organised regular meetings under the generic title of Operational Studies Seminars, and these have been well-attended and profitable to those taking part.

Nevertheless this was a closed and informal group: no formal records were ever published, and its discussions, being of the nature of progress reports, would not in any case have been particularly valuable to outsiders. The research team at Lancaster felt that the termination of its OSTI support in June 1969 afforded a suitable occasion for a wider discussion of this type of work, and its first independent activity was to organise a Seminar on Planning Library Services, with special emphasis on quantitative methods. A distinguished team of speakers was enlisted, and those taking part spent more than two days in discussing in some depth various aspects of the problem.

Proceedings were again informal: preprints were circulated by most of the speakers, who could then concentrate on highlighting specific points. This policy paid dividends, and the quality (and quantity) of the discussions was extremely high.

We have not attempted to reproduce discussions in extenso; and to keep production costs to a minimum we include in this volume the preprints exactly as submitted by the authors, with only four exceptions: Messrs. Brookes, Hindle, Leimkuhler and Vickery - for varying reasons - have revised their contributions for publication. All opinions quoted are those of the speakers and do not necessarily represent the views of the organisations to which they belong.

My thanks are due to all those who gave papers and took part in the discussions; to Mr. M. K. Buckland, who acted as Seminar Secretary and carried out all the complicated logistic exercises which made the occasion run smoothly; and to Mr. I. M. Stuart, who undertook the difficult task of reducing to a manageable form the many hours of spirited discussion.

A. Graham Mackenzie

SEMINAR MEMBERSHIP

(x Denotes part-time attendance)

x Dr. J.S. ANDREWS, University of Lancaster Library.
Mr. B.C. BROOKES, School of Librarianship and Archives, University College, London.
Mr. M.K. BUCKLAND, University of Lancaster Library.
Miss E.M. CAIRNS, University of Reading Library.
x Mr. C.F. CARTER, Vice-Chancellor, University of Lancaster.
Mr. H. COLHOUN, B.P. Research Centre, Sunbury-on-Thames.
Mr. H.F. DAMMERS, Woodstock Agricultural Research Centre, Sittingbourne.
Mr. A.M.E. DAVIES, Library, University College of Wales, Aberystwyth.
Mr. J.D. DEWS, Librarian, Manchester Business School.
Mr. R.M. DUCHESNE, Birmingham University Library.
Mr. R.A. FAIRTHORNE
x Mrs. M.J. FARRINGTON, University of Lancaster Library.
Mr. M.G. FORD
Mr. R.J. FULFORD, British Museum.
Miss H. GRAHAM, University Grants Committee.
Mr. L.C. GUY, School of Librarianship, Ealing.
x Dr. J. HAWGOOD, Director, Computer Unit, Durham University.
Mr. R. HEWITT, School of Librarianship, Birmingham.
x Dr. A. HINDLE, Department of Operational Research, University of Lancaster.
x Mr. J.L. ILLINGWORTH, University of Lancaster Library.
x Mr. D.F. JAMES, County Library, Kendal.
Professor F.F. LEIMKUHLER, Department of Industrial Engineering, Purdue University.
Mr. A. LONGWORTH, Librarian, Lancashire County Library.
Miss A.E. LUMB, University Library, University of Salford.
Mr. A.G. MACKENZIE, Librarian, University of Lancaster Library.
Mr. H.C. MACLACHLAN, Librarian, Paisley College of Technology.
x Mr. D.M. MAY, OSTI.
Mr. R. MORLEY, Department of Economics, Durham University.
Mr. W.E. MORRIS, Computer Unit, Durham University.
Miss J.G. O'CONNOR, Department of Astronomy, University of Leicester.
Dr. E.F.D. ROBERTS, Librarian, Trinity College, Dublin.
Dr. W.N. SCOTT, University Library, Queen's University of Belfast.
Miss C.A. SEYMOUR, Library Management Research Unit, Cambridge University Library.
x Professor R.N. SMART, Chairman of the Library Committee, University of Lancaster.
Mr. G.K.N. SMITH, Library Management Research Unit, Cambridge University Library.
Mr. W.H. SNAPE, Principal Lecturer-in-Charge, Liverpool Library School.
Miss E.A. STOW, Scientific Periodicals Library, Cambridge.
Mr. I.M. STUART, University of Lancaster Library.
x Mr. N. THOMPSON, Library and Information Systems Branch, D.E.S.
x Mrs. J.A. THORPE, University of Lancaster Library.
Mr. K. TIDSWELL, University of Bradford Library.
Dr. D.J. URQUHART, Director, National Lending Library.
Mr. B.C. VICKERY, Assistant Director (Research), Aslib.
x Mr. G.P.M. WALKER, University of Lancaster Library.
x Mr. I. WOODBURN, Department of Systems Engineering, University of Lancaster.
x Mr. B. WOODWARD, Science Library, Durham University.
x Mr. R. WORSLEY, County Library, Kendal.

SESSION 1: PROBLEMS OF LIBRARY PLANNING

H. F. Dammers	The library/information service in an industrial research establishment: planning and prospects
F. F. Leimkuhler	On information storage models
A. Longworth	Public library services
D. J. Urquhart	National libraries
Discussion	

THE LIBRARY/INFORMATION SERVICE IN AN INDUSTRIAL RESEARCH ESTABLISHMENT: PLANNING AND PROSPECTS

H. F. Dammers

1. INTRODUCTION

The increasing pace of change in our environment makes it ever more important to rationalise the manner in which we are trying to adapt the functions and operations of our organisations to changing demands. This planning activity is particularly important in an industrial research environment, which tends to be subject to changes in science and technology, as well as in commercial objectives. Such changes are likely to have an important bearing on the function and capabilities of the library/information service concerned.

Whereas for a university library the main problem might be stated as - how does on operate a given large system efficiently and adaptively?, in our case it appears to be - how does one develop system(s) that will meet demands in a rapdily changing environment?

The present paper attempts to discuss some of the problems associated with this type of exercise and this on the basis of experience gained in the development of an information service at the Shell Research Laboratories in Sittingbourne, Kent. The situation at Sittingbourne is perhaps of particular interest in that during the past 6-7 years the responsibility for developing and operating the information, as well as computer, facilities rested with one and the same group (Technical Information Services); this has led to a far greater degree of integration than is likely to be found in nearly any other organisation. It has obviously influenced very strongly the philosophy of approach. Many of the ideas in this area are still very fluid and may change with time.

2. OUTLINE OF DEVELOPMENT AT SITTINGBOURNE

In order to put our discussion in perspective, it may be useful to outline briefly the major developments relating to library/information services at Sittingbourne.

The formation of our library/information service dates back to 1955/56 when the European research activities of companies of the Royal Dutch/Shell Group concerned with its chemical products used in agriculture and public health were brought together in Shell Research Limited, Woodstock Agricultural Research Centre, at Sittingbourne; in 1959 this was followed by the establishment of the neighbouring Tunstall Laboratory, responsible for the study of environmental health aspects of Shell products. In 1963, the Milstead Laboratory of Chemical Enzymology was added to the research institutions on the site. As the name implies, this Laboratory was concerned with biochemistry-oriented basic research, dealing in particular with the study of natural products and micro-organisms. Although the creation of these establishments represents perhaps the major organisational change on the site, it is by no means the only significant one, as research objectives in various areas are continually subject to modification.

In a previous paper(1), an outline has been given of our actual and expected developments as regards mechanisation of our centralised information services. This, it was felt, could conveniently be presented as a series of development phases, of 2 years each, covering the period 1962-1972.

For the purpose of the present discussion, it appears more useful to use a breakdown according to the dominant type of information staff activity. This leads us to distinguish 3 stages of roughly 6-7 years each which may be briefly indicated as follows:

1955-61 - A period during which the main emphasis was placed on building up of adequate library facilities; the type of staff activity concerned was predominantly that of a technical librarian.

1961-67 - During this period the dominant role was that of the information specialists; coding, storage and retrieval of a variety of research data, the application of computer facilities for this purpose, and literature information searches were some of the major preoccupations.

1967-73? - This period starts with the availability on site of on-line computer facilities, allowing a much more efficient and intensive use of such facilities, and with the application of computer operated SDI systems. The emphasis appears to be more and more on the design and implementation of computer operated storage and retrieval systems adapted to specific site requirements. The dominant type of staff activity is perhaps that of the information technologist(1).

3. STAGE 1 (1955-61) - DEVELOPMENT OF LIBRARY FACILITIES

The main emphasis during this period was on the creation of proper library facilities, i.e. the provision of an adequate book collection, journal holdings and subject filing system.

A rather crucial point on which guidance was needed was how big the collection should be for the organisation served. Various routes were attempted in an effort to find an answer to this.

a) Plotting of the holdings of ca 30 industrial and scientific libraries, considered to be relevant to our case, provided a graph, presented in Figure 1(2). It relates the number of periodicals received with the total number of volumes held (books and bound periodicals). This graph provided us, inter alia, with a means of assessing whether the number of periodicals received was realistic in relation to the size of the library.

b) Another approach is to take data such as those published by Bourne(3) concerning the distribution of scientific and technical literature according to subject field and use this to relate the size of the library to those in other subject fields.

c) A third approach attempts to take into account the number of potential users. As a rule, one tends to find that the relation between the number of potential users and the size of the literature collection is not a simple linear one.

It has been suggested earlier(2) that the library, looked upon as the intermediary between the relevant world literature and the information already held by the users collectively, might perhaps, as regards size, approximate their geometric mean. (Figure 2) (Using the approach, it is difficult to see how the average British university library could justify a library of more than 200, 000 volumes - a view that is reinforced by cost effectiveness considerations)

Figure 1

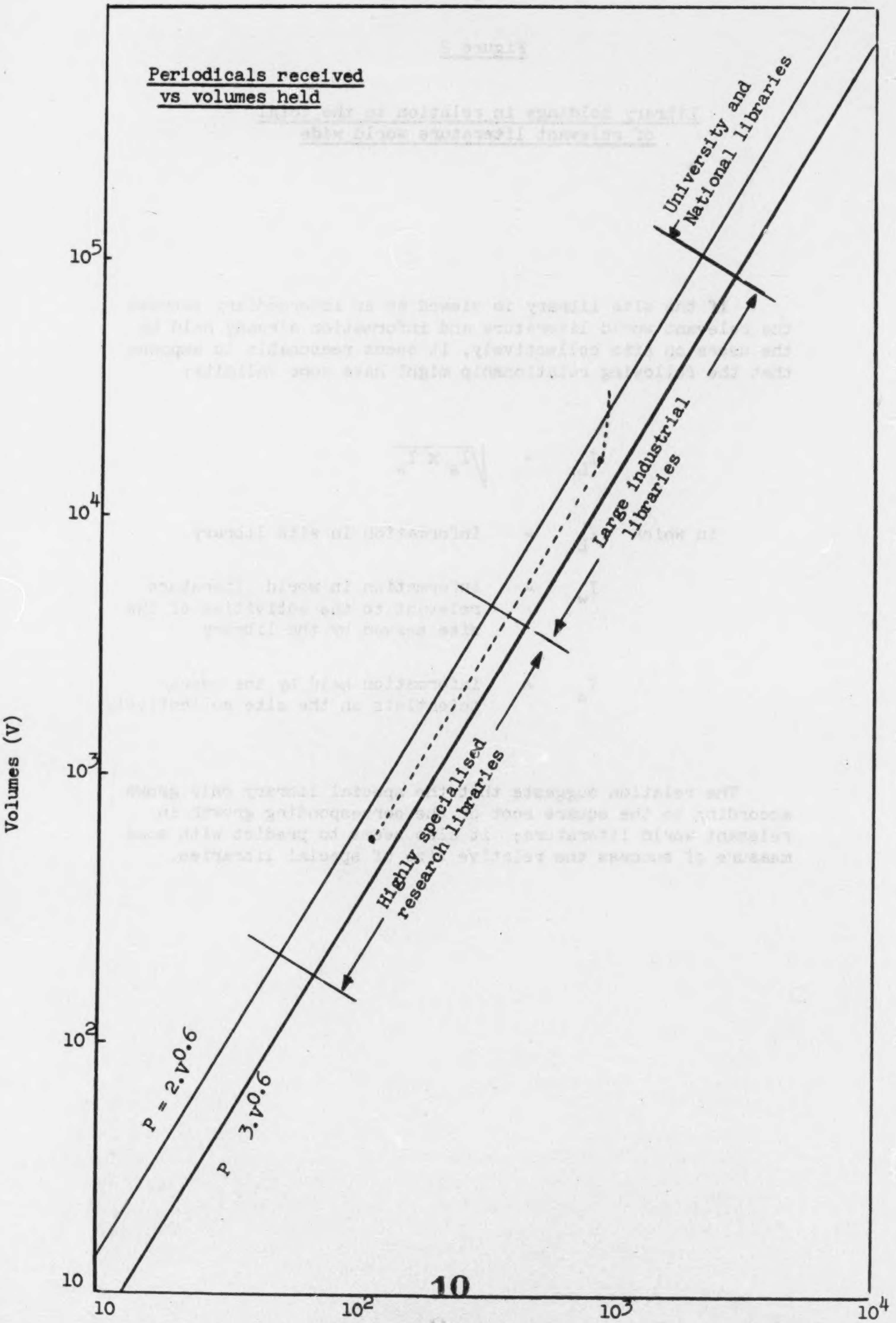

Periodicals received vs volumes held
University and National libraries
Large industrial libraries
Highly specialised research libraries
P = 2.V0.6
P 3.V0.6
Volumes (V)
10
10²
10³
10⁴
10⁵
Periodicals (P)
10
10²
10³
10⁴

Figure 2

Library holdings in relation to the total of relevant literature world wide

If the site library is viewed as an intermediary between the relevant world literature and information already held by the users on site collectively, it seems reasonable to suppose that the following relationship might have some validity:

$$I_L = \sqrt{I_s \times I_w}$$

in which I_L = information in site library

I_w = information in world literature relevant to the activities of the site served by the library

I_s = information held by the users/scientists on the site collectively

The relation suggests that the special library only grows according to the square root of the corresponding growth in relevant world literature; it also seems to predict with some measure of success the relative size of special libraries.

The various routes indicated are very dissimilar; hence if they all point to a similar result, one tends to accept the outcome with some confidence. In our case, they all appeared to be consistent with a library size of around 20, 000 volumes, which at the time appeared to be a quite acceptable estimate.

Another aspect during this phase that needs mentioning is subject filing. With the very limited staff available at the time, it was necessary to use a system for storage and retrieval of papers/documents of particular interest that would combine low labour effort in operation and maintenance with ease of adaptation and lack of 'predictive constraints'. This was the 'lattice index system' which has been briefly described in an earlier paper(4). It did indeed match the requirements it was designed to meet, i. e. it provided an effective subject filing system coupled with a highly adaptive indexing system that could follow changing research interests readily. The system is still in use although a major portion of its function has now been taken over by our computer operated systems. The system itself is now gradually being transferred into machine-readable form in order to make its relational information available for computer work on dictionary look-ups, associative searching etc. The lattice index made us aware of the cost of indexing systems. Even though this one, as far as subject filing/indexing systems go, was a relatively low cost one, it may well have absorbed at least £10, 000 in labour effort and materials by the time it had grown to a system containing ca 12, 000 files with the associated index.

This may sound substantial but is very modest when compared with various more elaborate indexing systems used in scientific laboratories, which may have cost £50, 000 or more, not to mention the basic indexing systems used in university libraries, the catalogues, which may represent an investment in labour and materials of £250, 000 - £500, 000. The awareness of the heavy investment in manual indexing systems, coupled with the difficulty of implementing modifications once they have grown to a substantial size, has led us to avoid developments in which such systems might have to play a major role.

4. STAGE 2 (1961-67) - MAIN EMPHASIS ON RESEARCH DATA HANDLING

This stage might perhaps be described as one in which the information specialists, as defined by Dolan(5), played a major role, i. e. most developments were those one might consider to be associated with this type of staff.

This arose first of all from the fact that far greater emphasis was placed on the handling of research data. During Stage 1, this had been dealt with by means of edge-punched card systems and later also by systems such as those described by Fletcher and Dubbs(6) using Kardex display systems. but by the early 1960's it became evident that a comprehensive system for our research data was required, and that only computer processing would offer hope of coping with the rapidly growing collection of data in a flexible and adaptive manner.

It was also clear that the comprehensive system needed to be developed centrally, i. e. not as a specific responsibility of one of the user groups concerned. Hence this task was undertaken by the Technical Information Services. It gave rise to a computer operated system for chemical structure storage and retrieval covering ca 50, 000 organic compounds: this system makes use of structure input in the form of IUPAC notation and on paper tape as produced by a chemical structure typewriter(7). It has already been dealt with at some length in earlier papers (8, 9), hence it needs no elaboration here. In addition, computer operated systems to deal with around 2 million biological test results obtained on these chemicals were developed; recording in this case has thus far been mainly on punched cards in fixed field format in order to allow handling as visual record, processing by punched card sorter and by computer(4, 10).

Activities such as these provided a challenging task area and training ground for our information specialists.

Moreover, in the first half of this stage there appeared to be a significant demand for information specialist type of activities such as literature searches and surveys, current awareness work in line with the then accepted function of the information specialist as a mediator between the body of scientific literature and the research worker. Dolan's definition(5) of the information specialists as being "those trained in a substantive technical field who have, in addition, some breadth of technical knowledge of the technical literature" appeared very appropriate to our case.

During Stage 2 we placed in fact considerable emphasis on the information specialist type of activity and the activities concerned did correspond quite well with those indicated in Figure 3, taken from Dolan's paper(5).

However, full implementation of the information specialist function to cover the main areas of research activities in our laboratories proved increasingly difficult to justify. The main reservations can perhaps be outlined as follows:

a) Various assessments suggested that really effective use of the information specialist approach in our case might require the number of staff of this type to be at least 5% of the number of research staff served. It was clear that it would be very difficult to justify expansion of our budget accordingly, partly because it is virtually impossible to demonstrate the economic merits of such an expansion.

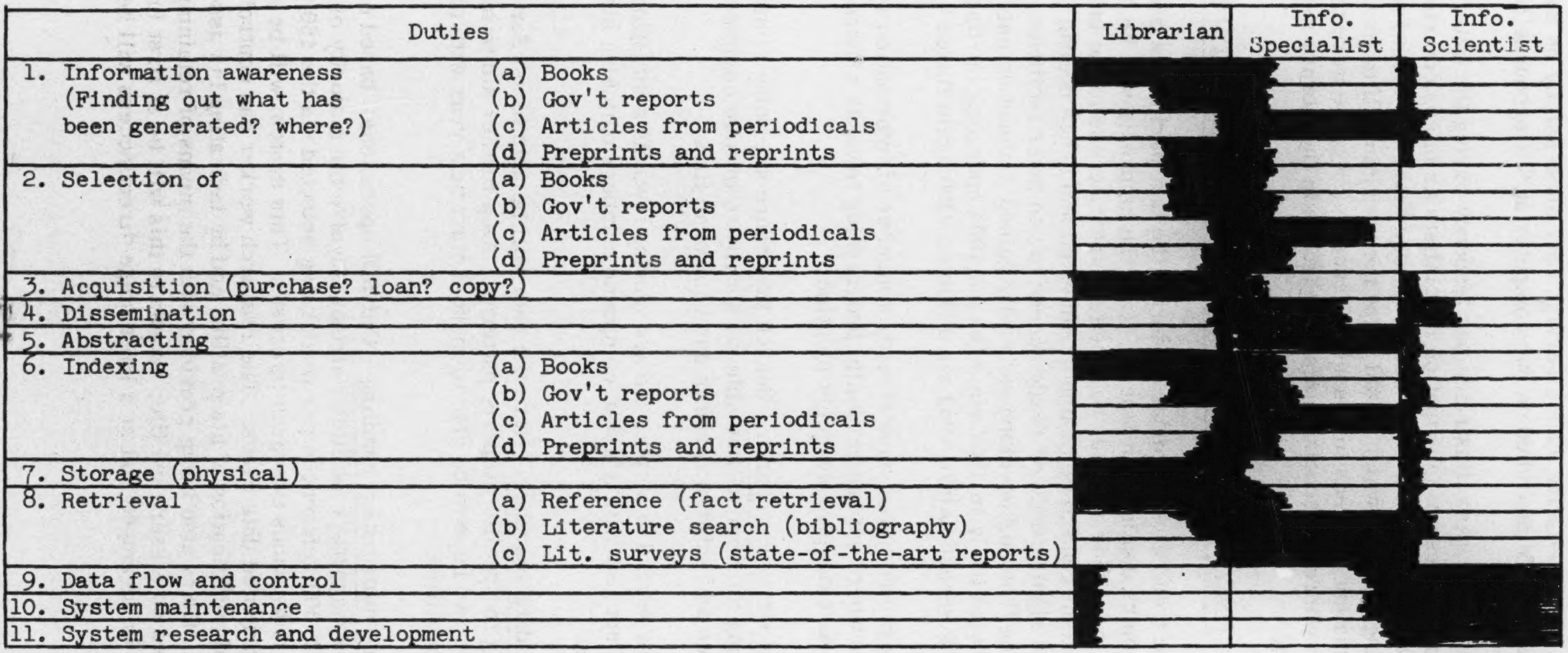
b) The quandary mentioned above is not made easier by the fact that, particularly in research, there is reason to doubt the merits of using an intermediary such as the information specialist in the information transfer process. This doubt appears to get some support from research such as that by Barhydt(11) which tends to indicate that the only reliable judge as regards the relevance of literature to a specific user requirement is the user himself.

c) Associated with this is the difficulty, pointed out in an earlier paper(2), of selecting in a current awareness service the papers considered to be of interest to particular groups of research workers. Selecting the 2,000 - 3,000 per annum obvious ones is relatively easy, but this is unlikely to satisfy those who are concerned with the more exploratory and innovating aspects of research work. Information from fringe territories is often of particular value in innovation and, in connection with this, it is of interest to consider the volume of literature at various possible levels of relevance depicted in Figure 4 and to ask oneself what level one should aim at when selecting literature for current awareness.

Considerations such as the above brought us to move towards systems which would enable the research worker to do the selection of relevant literature himself whilst relieving him from most of the drudgery associated with such a task. This was to be achieved by implementing a computer operated SDI system, whereby the user himself would be responsible for adjusting his search profile according to his requirement.

Figure 3

Work relationship between science librarian/technical information specialist/information scientist



Duties		Librarian	Info. Specialist	Info. Scientist
1. Information awareness (Finding out what has been generated? where?)	(a) Books			
	(b) Gov't reports			
	(c) Articles from periodicals			
	(d) Preprints and reprints			
2. Selection of	(a) Books			
	(b) Gov't reports			
	(c) Articles from periodicals			
	(d) Preprints and reprints			
3. Acquisition (purchase? loan? copy?)				
4. Dissemination				
5. Abstracting				
6. Indexing	(a) Books			
	(b) Gov't reports			
	(c) Articles from periodicals			
	(d) Preprints and reprints			
7. Storage (physical)				
8. Retrieval	(a) Reference (fact retrieval)			
	(b) Literature search (bibliography)			
	(c) Lit. surveys (state-of-the-art reports)			
9. Data flow and control				
10. System maintenance				
11. System research and development				

Acc. to Dolan, Int. J. Man. Machine Studies, 1, 39-50 (1969)

Systems of this type have been operated by us since early 1967(2, 12), using CT and CBAC tapes received from Chemical Abstracts Service. Since the beginning of 1969, we have been operating such a service very successfully on basis of CA Condensates tapes, which are searched on our computer at Sittingbourne(10, 13).

The above case illustrates our tendency during the latter part of Stage 2 to move away from our use of information specialists as intermediaries between the body of recorded information and the user; instead they have tended to become responsible for developing systems which would allow the user himself to interact direct with the body of information he needs to consult. Hence in a way we were tending to go back to the situation where the scholar himself searches in the library material for the information he needs.

5. STAGE 3 (1967-73?) - DOMINANCE OF THE INFORMATION TECHNOLOGIST

The trend already apparent during the last part of the previous stage is expected to become more dominant in Stage 3, i. e. the emphasis will be on the development of systems providing the user with better direct access to the information he requires. Such systems tend to be computer-oriented and hence depend for their development and operation significantly on ready access to computer facilities. It is therefore not surprising that this development really gained momentum only when we got direct access to computers early in 1967 via a Univac 1004 operating on-line to a Univac 1108, and during the second half of 1968 via a Univac 9300 again linked witha Univac 1108.

There are in our case already a number of information search and consultation activities where we are gradually progressing towards a more direct user/system interaction. As examples we might mention:

Chemical structure search - Search procedure are being simplified to such an extent that before long the user will be able personally to initiate computer searches for specific structures and substructures in our compound files.

With regard to output, we are moving towards a situation whereby the user will be provided not just with a list of compound numbers matching his enquiry but also structur formulae.

Property data searches - Here the user will be able to ask for tabulations; for searches according to specific property patterns. As a further aid to interactive use of the compu held files, we hope to be able to make a start next year with providing suitable visual display facilities.

Literature information handling - Our SDI operations, based on the use of Chemical Abstracts magnetic tape files, already involve the majority of research workers on our site; ca 2, 000 references per week being selected via the 150 profiles used at present to cover the various research interests. This system will be extended to include other biology-oriented data bases. The research worker can control the operation via the form lation and modification of his profiles and by indicating his assessment of the references selected. He is also being provided with the means of retaining selected and relevant references on a user tape file; access to this will be at first indirect, e. g. via a KWIC index; it is envisaged that at a later stage direct access will be provided, e. g. via a teletype.

Figure 4

Volumes of literature at various levels of relevance to research at Sittingbourne

		References/annum
1.	Literature references relating to specialist field	300 - 700
2.	Pesticides literature - highly relevant	2,000 - 3,000
3.	Pesticides literature - more general	10,000 - 15,000
4.	Pesticides Doc. Bulletin or Ringdoc.	30,000 - 50,000
5.	Bibl. of Agricult. or Chem. Titles etc	100,000 - 150,000
6.	All relevant fields (chem. and biol.)	300,000 - 500,000
7.	All scientific and technical literature	1 - 2 million

It is of interest to note that a surprisingly large proportion of systems appear to be working at level 2, whatever the field, suggesting that the level of selection tends to be subject-independent.

In addition, various other information and data files, relevant to specific users only, are being created in cooperation with the users and in other cases by the users themselves, with responsibility for maintenance and use vested in the user.

The information expert required for this type of development would appear to be the information technologist rather than the information scientist. In this we are following the definition given by Taylor(14) of information technology as being concerned with the development, design, and operation of information systems, whereas information science is concerned with the basic sciences underlying information system development. It would seem obvious that the information expert working in an operational environment such as an information service is more likely to function according to the former definition than to the latter.

The move towards direct user/system interaction poses demands as regards hardware and software which as yet are difficult to meet. It means the need for various direct access points (terminals) and this again generates the demand for processing and storage equipment capable of supporting the terminals in a flexible and efficient manner. The task of plotting an economically and operationally acceptable route for progress along these lines is a demanding one.

Such development also means that one can no longer ignore the integrated systems approach, i.e. in order to gain full advantage as regards economy and flexibility, one has to link related systems so that the costs of operations common to more than one subsystem can be spread over the system as a whole. This approach also makes it possible to incorporate greater flexibility and to provide more user options than is often feasible with separate systems. Such an approach, however, is as yet little explored and perhaps rather hazardous to implement.

Another trend which may well become more apparent during the next few years is that of the decline in coverage of the relevant literature which can be made available in one's own library as a result partly of the financial inability of libraries to keep up with the expansion of the world literature, and partly of the increased demand for current literature created by SDI services. This phenomenon has been discussed in an earlier paper(2); it may be of interest to illustrate it here with one of the tabulations given in that paper in 1967. (Figure 5)

The expectation, reflected in the table, was that our SDI service would stimulate interest in a wider range of literature sources and that the growth of our own site library would not be able to keep pace with this. As a result, borrowing from outside sources would increase strongly. The expected growth of outside borrowing given in the table actually corresponds with an increase of 50% per annum. The implementation of the SDI system using CA Condensates tapes on our site computer has led to a rather more rapid growth of this system than indicated in the figure; ca 150 profiles are now (mid-1969) being run. Analysis of the output suggests that the number of papers relevant to research at Sittingbourne is indeed of the order of 40, 000 per annum, but perhaps only 50% of these are considered significant. Borrowing from outside sources has increased by 50% comparing the first half of 1969 with the first half of 1968. In connection with this, it is of interest to note in Dr. Urquhart's paper that the National Lending Library has recently been experiencing an increase of ca 20% per annum in demands for loans. Although we will need to await the results of at least another year before we can confirm the trend with some confidence, it would appear that the results obtained so far are not inconsistent with the expectation expressed earlier, i.e. that the need for borrowing will

Figure 5

Estimate of literature relevant and borrowed at Shell Research, Sittingbourne

	1967	1972
Relevant papers in literature received	ca 15,000	ca 16,000
Relevant papers in literature not received on site	ca 10,000	ca 24,000 1)
Total	ca 25,000	ca 40,000
Relevant papers borrowed	ca 3,000	ca 24,000

1) i.e. 6 papers per 'profile' per week if 75 profiles are used in searching the literature tapes

continue to expand considerably and that therefore longer-term there ought to be a case for a national network allowing libraries to have direct access to large depositories of literature such as the National Lending Library.

This approach would envisage requests being sent via the computer network and the paper required to be transmitted for printing out on-site via a suitable on-line printing device.

The 1967 paper discusses in some detail the technical and operational feasibility of this approach. Developments since then have not given cause to change the view stated then that such an approach might well become viable during the first half of the 1970's.

In addition, we are likely to witness direct data transmission, e.g. between laboratories, replacing at least in part the present procedures based on transport of printed matter.

Trends such as those outlined above all appear to point to a decline of a variety of conventional information handling and transfer functions.

6. SOME OBSERVATIONS ON LONG TERM TRENDS

It may be obvious from the above that our situation with its very close link between information and computer services is as yet rather exceptional; furthermore, it is quite clear that one or two years of development along the lines indicated will be needed before one can confirm with some confidence the various trends postulated. Nevertheless, with these provisos in mind, one may perhaps be allowed to restate the overall trend which suggests itself on the basis of our experience.

It is that the information service will tend to become more and more a subsystem of the computer facilities serving the organisation concerned; at the same time, the responsibility for the actual operational use of information systems will more and more pass into the hands of the users themselves.

Again to illustrate this we may perhaps use one of the figures given in an earlier paper(1). This figure (Figure 6) did illustrate the total staff effort concerned with centralised information storage and retrieval and dissemination (ISRD) activities at Sittingbourne expressed as a percentage of the total staff on site for the years 1962-72.

During the 2 years since 1967, the relative decline in staff effort has been much faster than foreseen in Figure 6; this decline is associated with a much more intensive use of computer facilities made possible by the provision, since May 1967, of direct access to computer facilities at Sittingbourne.

The decline is accompanied by a move of some of the staff towards more general systems work on computer applications. The way things are developing in our case suggests that this decline is going to continue, if not at the same rate as experienced in the period 1967-69.

It is very hazardous to predict longer-term trends in our rapidly changing environme
Nevertheless, our experience so far appears to point at the possibility that in our case the centralised ISRD facilities may perhaps gradually disappear as a separate operational

Figure 6

Staff effort associated with centralised information storage, retrieval and dissemination activities

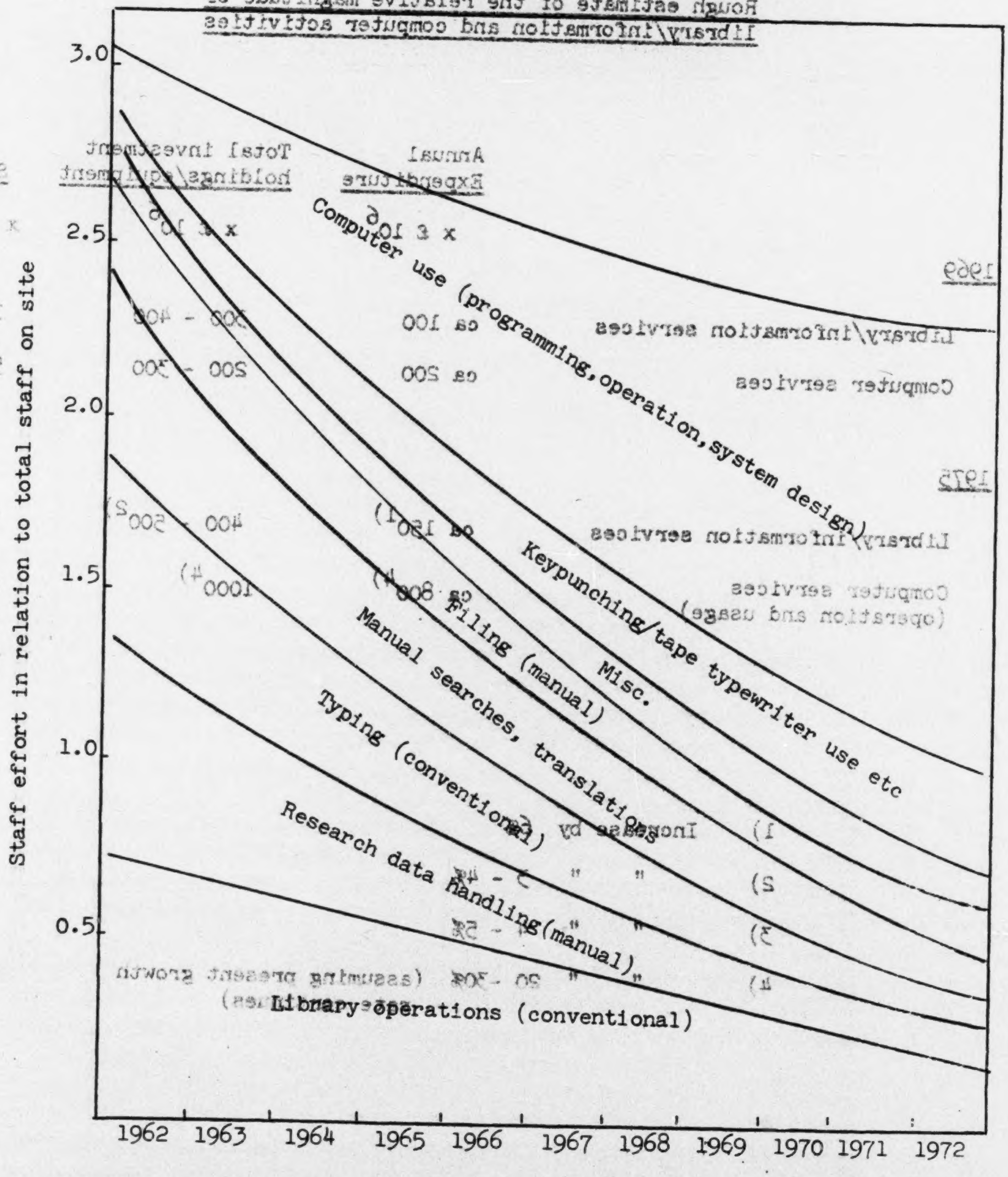


Figure 7

Rough estimate of the relative magnitude of library/information and computer activities

	Annual Expenditure	Total investment holdings/equipment	Sta
1969	x £ 10^6	x £ 10^6	x 10
Library/information services	ca 100	300 - 400	10-
Computer services	ca 200	200 - 300	50
1975			
Library/information services	ca 150 1)	400 - 500 2)	15-2
Computer services (operation and usage)	ca 800 4)	1000 4)	200 4)

1) Increase by 6%

2) " " 3 - 4%

3) " " 4 - 5%

4) " " 20 -30% (assuming present growth rate continues)

entity during the next decade. The user would feed his data/information in direct and consult the stored information direct; a central information service would no longer be operationally involved in this process. The task of the information expert would primarily be concerned with information system design and development matters and no longer with operation.

Whether such a development is likely might, in part, be judged from the projected manpower, and capital resources devoted to library/information work on the one hand, and computer activities on the other. (see Figure 7)

Perhaps we may also see the gradual decline of medium-sized specialist libraries such as ours, although this would depend on the inherently slow development towards effective large regional and national data and information centres as part of a national network. The future task of the information expert remains a very considerable one; however, it will be very different from the one he has conventionally been associated with.

REFERENCES

1. Dammers, H. F. (1968)
Phased transition from non-mechanised information storage and retrieval to on-line computer operation
Proceedings F. I. D. /I. F. I. P. Conference on Mechanised Information Storage, Retrieval and Dissemination, Rome, June 14-17, 1967, 415-436
North Holland Publishing Company, Amsterdam

2. Dammers, H. F. (1968)
Integrated information processing and the case for a national network
Inform. Stor. Retr., 4, 113-131

3. Bourne, C. P. (1962)
The world's technical journal literature: An estimate of volume, origin, language, field, indexing and abstracting
Am. Doc., 13, 159-168

4. Dammers, H. F. (1966)
Computer handling of literature information and research data in an industrial establishment
Proceedings 36th International Congress on Industrial Chemistry, Brussels, 1967

5. Dolan, F.T. (1969)
The role of the information scientist
Int. J. Man. Machine Studies, L, 39-50

6. Fletcher, J.H. and Dubbs, D. S. (1956)
Quick access to research records
Chem. Eng. News, 34, 5888-5891

7. Mullen, J. M. (1967)
Atom-by-atom typewriter input for computerised storage and retrieval of chemical structures
J. Chem. Doc., 7 88-93

8. Dammers, H. F. and Polton, D.J. (1968)
Use of the IUPAC notation in computer processing of information on chemical structures
J. Chem. Doc., 8, 150-160

9. Polton, D. J. (1969)
Conversion of the IUPAC notation into a form for computer processing
Inform. Stor. Retr., 5, 7-25

10. Dammers, H. F. and Gallagher, P. J. (1969)
Computer use in information and data handling
Automation, Mechanisation and Data Handling in Microbiology
Society for Applied Bacteriology Technical Series No. 4
Academic Press (in press)

11. Barhydt, G.C. (1967)
The effectiveness of non-user relevance assessments
J. Doc., 33, 146-149

12. Gallagher, P. J. (1968)
Some experiences with selective dissemination of information
Information Scientist, 2, 103-106

13. Gallagher, P.J. (1969)
Development and operation of an SDI system using CA Condensates tapes
Unpublished communication WP-231-69
Shell Research Limited, Woodstock Agricultural Research Centre, Sittingbourne, Kent.

14. Taylor, R.S. (1966)
Professional aspects of information science and technology
Annual Review of Information Science and Technology, 1, 15-40

ON INFORMATION STORAGE MODELS [*]

by

F. F. Leimkuhler [**]
Purdue University

Information storage theory is not a well-defined area of research in the formal sense and one is still free to make of it what he wants. My own viewpoint is that of an industrial engineer and operations researcher who has been seeking ways to develop mathematical models for the analysis and design of library-type information storage systems. The word "storage" can evoke some bad vibrations in library circles where it is associated with the least respectable aspects of librarianship, but I choose to use the word in a broad and inclusive sense. While it may be more meaningful to define libraries as communications systems which transfer information through time and space, such transfers are accompanied by significant delays which give rise to meaningful storage problems. The study of libraries from the storage viewpoint can identify some crucial aspects of information systems which are often overlooked or ignored when the focus is on communication.

---

[*] Prepared for a Seminar on Planning Library Services, University of Lancaster, England, July 9-11, 1969. The first part of this paper was read at the Sixth Annual Information Retrieval Colloquium, Philadelphia, May 9, 1969.

[**] Visiting Professor of the University of California, Berkeley, during 1968-69. This research was supported in part by National Science Foundation Grant GN-759.

Although I have been working at storage models for several years and have advocated its practical importance to libraries, I am quite aware that it is not an easy matter to translate theory into practice. It has been my experience that action follows from need, but that better action can result if some good theory is available to help diagnose the need and to guide the remedial efforts. Our present theory is quite rudimentary from a research viewpoint and has a good way to go before it reaches the sophisticated state of, say, modern inventory theory, but the beginnings are there and there is good promise of a rich harvest.

It seems a bit ironic to me that the first important exploitations of this new study of library systems and the best guarantee of its continued support are probably going to come from outside the library proper. It is in the design of automated special purpose information systems that one has the greatest freedom to make innovations and the greatest pressure to apply a uniform economic and technical yardstick to every facet of the system. Conventional libraries are already mature technological systems in their own right. They are predicated on an earlier piece of mechanical wizardry -- the printed book and a well-developed clerical work system to support its exploitation. Within these bounds most of the waste has been trimmed away in the long lean years of experience. As with a magnificent old clock, one doesn't tamper with it. If you want to keep it running, you will have to find parts made to the original specifications.

Still, I find it fruitful and hopeful to pursue the study of information storage systems in the context of conventual libraries. They offer a rich source of experimental data, and a wealth of ingenuity in an operational setting. As a "going" system, it is a good place to test the validity of one's models. Furthermore, there are plenty of indications that some radical changes in library operations are going to have to occur in the not too distant future.

Space Models

A good first example of the nature and implications of storage theory is the book shelving model which was developed at Purdue several years ago (1, 2). The model assigns a given collection of books to a set of shelves with those lengths and heights which will minimize the shelf area required. The direct application of this model to some representative library collections has indicated that relatively efficient storage can be achieved by using only three or four different heights; and, in fact, by shelving large books on their fore edge, one can do remarkably well with only two shelf heights (3). This result poses an interesting question to those large libraries which presently employ eight or more size classifications in their depository-type storage areas. It also calls into question the wisdom of spending extra money on variable height shelving and the practice of adjusting shelves up and down as new books are added. But these are rather minor benefits; and, perhaps, the greatest immediate importance of the model is its ability to show rigorously that one is not going to achieve dramatic reductions in space utilization through shelf arrangement alone. If all books could be stored by size on their fore edge, the best one could do is to double shelving capacity (4). This is not by any means a long run solution to library storage problems.

Of considerable interest are two recent applications which are peripheral to the shelving problem. In one instance, the MARC catalog tapes produced by the Library of Congress were examined for the distribution of lengths of the records they contain (5). Some 65 different record lengths were found in a random sequence. When ordered by size, they formed a bell-shaped distribution. The book shelving model was applied to find the optimal record lengths to use for blocking the tape so as to produce a fixed record length tape with one, two, three, etc., different block sizes. The records would be in conventional sequence within each block size, and shorter records would cause some loss of storage space. As with books, it was found that the use of only a few block sizes could make fixed length processing relatively efficient at the expense of more storage capacity. Again, as with flipping large books on their fore edge, the model could be used in conjunction with a program for selective code compaction of longer records so as to achieve an optimal balance of processing and storage costs.

An even more esoteric application of the basic shelving model is the possibility of using it in the production of microform records. For example, in producing microforms of conventional book material, one has to compensate for the variable sizes of book pages. What set of fixed frame sizes would achieve an optimal balance between the cost of handling variable frame sizes and the cost of reproducing blank spaces? If a single frame is used, it must accommodate the largest page size at the expense of much excess capacity for smaller pages. If only two sizes are used, what smaller size is optimal? The use of more sizes decreases lost area but increase the complexity of the system. Again, as with the fore edge storage of books and the compression coding of MARC tapes, variable magnification might be used in conjunction with the selection of optimal frame sizes. This complicates the analysis considerably but allows for many more options and the possibility of a much better solution.

Another different sort of application of the book shelving model was made recently to the design of industrial warehouses, where the problem was that of determining optimal bay configurations and the assignment of variable size lots of palletized materials (6). This might have useful implications for the design of library building and the assignment of subject groupings of varying size to different areas so as to minimize the sum of paging and space costs.

Usage and Cost Models

Space models of the kind considered above have the analytic advantage of dealing with the physical measurement of inanimate objects and avoiding the more difficult problems of measuring human behavior and judgment. It is the absence of the human element which makes them most amenable to mechanical applications and which evokes the strongest suspicions of practical librarians. There are two basic ways of approaching the role of human intervention in man-machine systems. One way is to take the direct approach and concentrate on people, their perceptions and reactions to the system. This is the approach of the behavioral scientist.

A second approach is an indirect one of focusing on the physical components and attributing to them attributes which are really the net effect of some prior human action. For example, we speak of a book circulating, of it containing certain information, or of it having so much worth and relevance. This approach permits the reduction of much of the human element to measurable quantities which can be related directly to other aspects of a system. This is the approach of the economist who can infer a value measurement from the limited availability of certain resources and the desire to have more of everything rather than less.

A good example of this approach is in the work of Philip Morse and his recent book on Library Effectiveness (7). Most of his models depend heavily on the notion of "randomness" in the behavior of library patrons. Tossing coins to retrieve information is an idea which seems patently absurd, if applied to some individual researcher, but is remarkably useful in measuring the collective effects of many individual choices and actions on the performance of a service system. Once we accept these measures as good approximations, we are in a position to make meaningful comparisons and recommendations for system improvement.

The analysis of depository schemes for libraries is a good example of what I call usage models. In general, depository models have argued that a considerable portion of a library's collection is so rarely used that these items could be stored elsewhere at less cost or to make room for new material. On the basis of out-of-pocket library costs alone, Lister (8) argued that several science libraries at Purdue could justify the storage of up to 60% of their holdings and achieve a small reduction in total costs. However, if some significant user delay cost is added to the charges, the advantages of depository storage are reduced drastically. The net effect of his study is to show that depositories do *not* provide an easy solution to library storage problems. Where space is limited and storage is the only answer, however, Lister's models do show how a rational, suboptimal policy can be developed.

A variation on this theme is seen in the recent study by the Center for Research Libraries (9) of the feasibility and potential benefits of a cooperative storage and lending facility for periodicals. This study is notable for its analysis of pertinent cost data from several libraries. A similar preliminary study was made in England at the University of Lancaster which showed that university libraries might utilize a national lending service for 10 to 30% of their demands depending on the user delay cost.

Perhaps a better prototype storage model of the usage variety is the one proposed by Cole (10) and refined and extended by Buckland (11). Cole showed that a 2000 volume petroleum library could expect to satisfy the greatest number of user requests by subscribing to approximately 190 journals and holding them for about 11 years. He assumed exponential obsolescence of older volumes and a Zipf-type pattern for the marginal productivity of additional journal titles. Buckland introduced some considerable mathematical refinement to these basic relationships and was able to go beyond Cole's results and show how to include such additional features as variable retention periods for different journals and use of interlibrary loan options. He also looked at how to meet a given level of service and minimizing a cost function that gives explicit recognition to different storage policies. An excellent review of the history of library use studies and models was made by A. K. Jain (12).

Somehow the usage models, like the space models, seem to fall short of the mark in an attempt to come to grips with the critical problems of libraries. Libraries do have something in common with warehouses and bookstores but there is still a residual difference which cannot be ignored. The further development of economic models of library-type systems must focus on investment as well as operational costs. Because of their patterns of long-term storage and exponential growth, investment models may provide the better approach to the understanding of library economics. This approach would seem to be better suited to the development of system planning models and the justification of technical innovations.

Retrieval Models

One cannot pursue information storage models very far without confronting difficult problems of information retrieval. It is interesting to observe how operations researchers and industrial engineers have tended to focus on the storage side of library systems while library scientists focus on the retrieval side. M.E. Maron has defined "the library problem" as the problem of retrieval and not of storage. He points out that the space problem is largely a problem of technique and economics -- a matter of miniaturization, for which the necessary physical theory is already available, but that any use of miniaturization or mechanical storage is going to necessitate the development of a sophisticated remote access capability. The theoretical work on information retrieval which is necessary for such a development is not available now.

The separation of storage and retrieval is the critical factor in the automation of information systems and libraries. Conventional libraries must depend on direct user access to keep costs within bounds and to make card catalog systems work, i.e., we really have catalog-aided manual retrieval. Interlibrary loans, for example, are one of the most expensive kinds of services a library offers; and yet even this is cheap when compared to the cost of providing remote reference service, as for example, in the specialized information centers which the government has funded.

A thorough review of retrieval models is too large a subject to cover here and is beyond my competence to review. There appears to be a wide variety of approaches and classes of models, among which are those based on behavioristic studies of how man uses language in the transfer of information; and then, there are the computer-oriented approaches which concentrate on the algebraic and physical capabilities of electronic devices.

The approach I have taken is an operational one in that it attempts to model the patterns observed in existing working systems, and to draw inferences for local optimization and evolutionary development. This is the method of operations research as opposed to basic research and is not offered as a substitute for the latter but as a complimentary approach.

It is characteristic of OR work to look for analogies from other fields and to draw heavily on the selective experiences of past observers so as to attempt a sort restatement of what is known about a system in the language of applied mathematics. An example is the model which appeared recently in American Documentation (13) where ideas were taken from the theory of military search and reconnaissance and from the earlier empirical work and wisdom of the English documentalist, S. C. Bradford. These ideas were used to formulate an analytic model that also incorporates a mathematical approach which is similar to the math used in the book shelving model, and which, by the way, is developing its own separate history within OR circles under such names as the assignment problem, the cutting stock problem, the packaging problem, and other such titles.

This approach seems promising, expecially in the connection it made with Bradford's "scattering" studies, which can be related to Ziph's law and which, in turn, opens the door to some promising extensions into information theory, linguistics, and economic theory. Furthermore, the first model has called into question the proper measurement of search effectiveness and its relation to user preferences, perceptions, and behavior (14), and the relation of the latter to expect judgment of the relevance and content specification of a chunk of printed matter we call information. Thus, the models have a double payoff: they can lead to practical applications and can also open doors to new theory.

References

(1) Leimkuhler, F. F., and Cox, J. G., "Compact Book Storage in Libraries", Operations Research, 12:3, pp. 419-427, May-June, 1964.

(2) Cox, J. G., "Optimum Storage of Library Material", Ph.D. Dissertation, Purdue University, June, 1964. See Reviews in Library Quarterly, 35:3, College and Research Libraries, 26:3, and Library Review, 20:2.

(3) Popovich, J. D., "Compact Book Storage", M.S.I.E. Project, Purdue University, 1966.

(4) Raffel, L. J., "Compact Book Storage Models", M.S.I.E. Thesis, Purdue University, June, 1963.

(5) Stirling, K., Cost Exchange Analysis of Variable Length versus Fixed Length Marc II Bibliographic Records, Course 244-C, School of Librarianship, University of California, Berkeley, Fall, 1968.

(6) Roberts, S. D., "Warehouse Size and Design", Ph.D., 1968, Purdue University, Major Prof. Ruddell Reed, Jr.

(7) Morse, P. M., Library Effectiveness, The M.I.T. Press, 1968.

(8) Lister, W. C., "Least Cost Decision Rules for the Selection of Library Materials for Compact Storage", Ph.D. Thesis, Purdue University, January, 1967. U.S. Clearinghouse Report PB 174441.

(9) Williams, G., "Library Cost Models: Owning Versus Borrowing Serial Publications", Center for Research Libraries, Chicago, 1968.

(10) Cole, P. F., "A New Look at Reference Scattering", J. of Documentation, Vol. 18, No. 2, June 1962.

(11) Buckland, M. K., and Woodburn, I., "Some Implications for Library Management of Scattering and Obsolescence," University of Lancaster Library, Occasional Papers No. 1, 1968.

(12) Jain, A. K., "A Statistical Study of Book Use," Ph.D. Thesis, Purdue University, January, 1968, U.S. Clearinghouse Report PB 176525.

(13) Leimkuhler, F. F., "A Literature Search and File Organization Model", American Documentation, 19:2, pp.131-136, April, 1968.

(14) Baker, N.R., "Optimal User Search Sequences and Implications for Information Systems Operation", School of Industrial Engineering and The University Libraries, Purdue University, 1968.

PROBLEMS OF LIBRARY PLANNING

PUBLIC LIBRARY SERVICES

By

Alan Longworth, County Librarian,
Lancashire

In presenting a paper such as this, a paper which is one of fifteen in a seminar concerned with a very precise topic, it is specially necessary to define one's topic and terms of reference if undue duplication is to be avoided. For the next few minutes therefore I shall be speaking of the public library as it has been understood in Britain, and through many other parts of the world, since the passing of the first Public Libraries Act in England in 1850, and as it is still understood, with some variations, under the terms of the current Public Libraries and Museum Act of 1964.

A "public" library is one provided by a local authority under the terms of appropriate Local Government Acts, and Public Libraries Acts, and primarily financed from local rates.

I could almost define my topic, even today, by quoting the preamble and terms of the Act of 1850 but I will only refer specifically to the clause which states that the purpose of the public library is to provide books for education, information and recreation, freely available to all who wish to use them.

In the above sentence I have, almost unwittingly, identified the differences between "public" libraries and the other three categories of libraries which are being discussed this afternoon.

(a) Public libraries are authorised and governed by statute.

(b) Public libraries are for use by the public at large.

(c) Public libraries are financed from public funds raised at local level.

(d) Public libraries are expected to provide education, information _and_ recreation.

It is in consideration of these four points that the main differences in the four types of libraries, and the special problems of planning a public library service will emerge.

1. Statutory Provision and Control

It is particularly appropriate that a seminar, meeting in Lancaster in the summer of 1969, should hear a paper on library planning which is presented by the County Librarian of Lancashire, who was formerly City Librarian of Salford. Truly Lancashire can claim a major share in the pioneering of the public library movement, and the rivalry between Warrington and Salford for the honour of opening the first free library is perhaps weighted towards Salford by the work of Joseph Brotherton in introducing the 1850 Act into the House. In contemporary terms Lancashire County Library faces dissolution and the area which is now called Lancashire faces a literally complete re-organisation of its local government structure if the recommendations of the second Maud Commission are accepted and promulgated. Whatever the future of the report, Lancashire will be significantly changed by the creation of the Central Lancashire New Town and Warrington New Town.

The statutory provision of public libraries is probably the most important single difference between "public" and other libraries and must be clearly understood. The local authority is not a law-making body and can only operate under powers bestowed on it by Central Government and within limits very clearly defined. In 1969, these are the 1964 Public Libraries and Museums Act, and various Local Government Acts, of which that of 1933 is possibly the most important. The possible 1970 Local Government Act, based on the Maud report will stipulate the boundaries of the new authorities, the composition of their governing bodies, the powers they are to have and the methods they will use to finance the work they are to do.

This, complemented by the 1964 Libraries Act, will provide the framework within which the public library service of the future will be planned.

2. The Public-at-Large

It is in the category of library user that there is the greatest practical difference in our four types of library. The special library has an extremely limited clientele both in numbers and in range of library interests. The university library will probably have larger numbers to deal with and also a wider range of topics, though university students and staff, are by the nature of their selection, rather special categories of people.

National libraries, or some of them, are, theoretically, for use by the general public. In practice this is not really true, since carefully vetted introductions are needed for the British Museum whilst the National Lending Library is extremely careful even in its choice of agents.

It goes without saying that the users of the public library are "the public" and by "the public" we really mean the whole of mankind - young and old, black and white, deaf, blind, insane, uneducated, uneducable, genius, autistic - every conceivable category into which people may fall.

But once we set ourselves the problem of "planning public library services" in the context of this conference we immediately limit our understanding of this word "public". We limit ourselves to a public which is English-speaking, to people who understand the words we speak and the signs we write and print.

Because we are planning a library service for "the public", in Britain, in 1969 - let us think about this "public" in some little detail. It would be extremely useful if we could create a single personality who could embody all our public's characteristics in one person, a sort of contemporary John Bull. Physically this might be almost possible, but we are not concerned with the physical body, we are concerned with men's minds, and variations in mental capacities and characteristics are immensely more variable than physical make-up. We must therefore allow for all the mental attributes which go to make up the individual personality.

Another aspect of man's mind we have to consider is its state of training and education. We must cater for everyone throughout the whole educational range from the barely educable backward reader to the most highly trained brain our educational system can produce. We must deal with students at all stages from infants' departments to primary, secondary and further formal education, the particular requirements of modern vocational training and re-training, and the less purposeful, but none-the-less personally staisfying activities of non-vocational training in the way of hobbies.

We must also consider the widely varying library needs of people throughout the whole of their lives and this is worth some detached consideration.

Babies obviously do not read, though from quite an early age they do begin to understand the significance of visual symbols. Children normally learn to read between the ages of four and seven years, though backward readers may be some years late and in a few cases may never really read. The normal child discovers the magic of imaginative reading at seven years of age and this phase lasts about four or five years. During these years children will read their own comics, mum's magazines, dad's newspapers and every book, "suitable" or "not suitable", they can get their hands on. During these years an avid reader will get through a book a day as well as a considerable amount of reading of other material.

At the age of twelve or thirteen the magic seems to disappear, and this change with the onset of puberty is a subject for psychological research which could have enormous significance. Whatever the reason, the magic disappears and reading seems to become more purposeful. In grammar schools and higher education this is undoubtedly connected with syllabus study but in modern schools where the academic pressure is not so great the magic still seems to disappear and reading loses its charm. In all forms of advanced formal education, private reading, along with supervised experimental work, lectures and tutorials, takes up all the student's time - but here I begin to stray on to other fields to be covered by other speakers.

When formal education is over the continuation of regular reading depends to a considerable extent on a persons basic intelligence, former education and kind of job. The professional man reads for both work and pleasure throughout the rest of his life. His wife who normally has a broadly similar background reads mainly for pleasure and relaxation during the early years of family life, her reading growing broader and more purposeful as the family grows up and she has more time and energy to follow her own pursuits.

In working class homes reading is largely confined to newspapers, magazines and paperbacks of the more popular and less mentally demanding kind.

In later life when the husband has retired from working life, reading helps to fill his days and use of the public library is frequently one of the regular features of his more settled life, and this more leisurely pace of his life also seems to affect his wife who may also return to reading to fill part of her spare time.

What I have just described is a fair cross-section of readers who use the public library - children, students, young marrieds, professional people, and older people. It is interesting to speculate on the motives which bring them to the library, and it is also interesting to speculate on the reasons why other people do not use the library.

In this context of the public-at-large it is also important to identify the size and character of the communites in which people live. Size can range from the individual country cottage and farm, the isolated hamlet, the small country town and the industrial conurbation, to the provincial or national metropolis. Basic standards are now laid down by the Department of Education and Science, suggesting the size of buildings and kind of service which needs to be provided for each size of community, but the Department recognizes that the character of the community also needs to be considered. For example in Lancashire, Grange-over-Sands is a small seaside town of little over 3,000 population, of whom 85% are retired business and professional people, needing a library service but needing very little help and guidance in its use. At the other end of the scale is Kirkby, housing 60,000 of Liverpool's overspill population - 75 % of them under 21, posing considerable problems of community control and discipline. Here help and guidance in using the library is of paramount importance and the crusading work of utterly dedicated senior librarians and staff is an object lesson in community service.

3. Finance

Unlike the special libraries, but like university and national libraries, the financing of public libraries is from public funds; but unlike the other two, public library funds are under local rather than national control. The Department of Education and Science now lays down certain standards of expenditure and "encourages" local authorities to achieve these standards. The Department also controls the amount of expenditure on buildings by its authority to limit the use of capital funds, whilst the annual revenue spending of local authorities is supported where necessary by grants from Central Government funds and part of the rate support grant finds its way into the library budget.

Unfortunately, in the world of local government finance, the value of the library service is not always recognized, and the library always has to bid for its budget in competition with financial needs for housing, health and education, and other major and minor services. In times of national and local financial crises, the less obvious needs of the library are frequently sacrificed in order to maintain the funds of a service whose value is more immediately apparent. I am quite sure that the controls on expenditure in all libraries are extremely stringent and equally irksome.

4. The Supply of Books and Other Materials

Because its clientele is the public-at-large, and because its terms of reference are to supply books for education, information and recreation, the public library's buying policy is completely different from the policies of the other three bypes of library. In this it is to some extent the victim of its own history. In nineteenth century public libraries one of the major motives was to encourage members to read, to read anything at all, and the recreational aspect of library work was of considerable importance in the days before radio and the visual entertainment of the cinema and television. For this reason the public library came to be regarded to a considerable degree as the source of pastime reading for the less wealthy and this image is still unfortunately not entirely erased. Since the advent of other forms of domestic entertainments, and the consequent diminishing of the demand for purely recreational reading, public libraries have been able to concentrate their resources more on providing books and other material in the fields of information and education, and in a world in which life daily becomes more complex, and in which education continues to develop at a staggering pace, the use of the public library continues to grow.

In the last few years, the Central Government has begun to encourage the use of public libraries as general cultural centres rather than simply as book repositories and in small and medium sized communities the library is now a meeting place, and the recognized venue for cultural exhibitions and displays of all kinds.

In terms of the physical presentation of material, public libraries tend to remain faithful to the book, and though collections of micro-documents, tapes, films, gramophone records and other forms are not unknown, their impact is still slight.

Rightly or wrongly, the public library's success is still to some extent measured by the number of books issued, and the value of books which are not in constant use may be questioned. Certainly in small and branch libraries, each book needs to justify its shelf space in terms of number of loans. This also leads to heavy wear-and-tear and the importance of specialised channels of supply for reinforcing books, from specialised public library suppliers.

The decline in sales of "hardback" books in bookshops, the decline in "commercial" libraries in the past twenty years and the development of the paper-back book trade has meant that the public library market is now a much more important factor in the mind of the publisher when he decides to publish a book, but that perhaps is leading us outside the field of our present argument, though without publishers there could be no libraries at all.

5. Planning

In the preceding paragraphs I have discussed the factors which affect the planning of a public library service and in particular those factors which are peculiar to public libraries and the differences which exist between public libraries and the other kinds of libraries.

Once these factors are fully understood, the actual planning of the library can proceed along the traditional lines, of books, buildings and staff; control, administration and systems; and finance.

If the report of the Royal Commission on Local Government is accepted, the public library service, like local government itself, will have the opportunity to re-assess its purpose and methods, and of re-organizing itself so that it can continue to play a proper part in man's continuing evolution in the latter part of this century and the beginning of the next.

PLANNING LIBRARY SERVICES - NATIONAL LIBRARIES

by D.J. URQUHART

I think it would be wisest in the present circumstances if I confined the majority of my observations to my experience at the N.L.L. The case for creating the N.L.L. and deciding what sort of library was required depended upon a number of measurements. The essence of these has been set down elsewhere.(1) Looking back it seems to me that the data was relatively simple to collect and analyse. The more difficult part was to decide what might be relevant. Such decisions were guided by a simple philosophy that it was important to "collect facts, analyse facts and act on facts and ignore opinions".(2)

This paper will endeavour to consider how far this philosophy of observation and measurement has been used in operating the N.L.L., how far the decisions have depended upon intuition, and how far the experience is relevant to other libraries. As I start to write this paper let me say that its purpose is exploratory for I do not know what the answers will be.

To help me to focus my attention on the problem I will first try and set down the main changes which have occurred since the N.L.L. was officially opened on the 5th November, 1962. Since then the scope of the library has been expanded to include some parts of the medical field previously excluded, and more recently to serials in all parts of the social science field. The number of current periodicals received has risen from 14,700 in April 1963 to 32,500 in April 1969. The number of requests satisfied by the library has increased from 187,600 in 1963/4 to 651,800 in 1968/9. In the last twelve months the percentage of photocopy requests has about doubled. In addition, a variety of courses on the use of scientific and social science literature are now being run and the N.L.L. is responsible for the U.K. input into MEDLARS and the final formulation of MEDLARS searches.

The main restraint on the N.L.L.'s development has been the limitation of staff numbers. To cope with this and the recent volume of activities there has been a number of procedural changes. The overall result has been that despite the rising cost of literature and rising salaries and despite the increase in "non-productive" work such as MEDLARS, the unit cost per issue in pounds has fallen - see Table 1 below. Nevertheless the N.L.L. success factor has risen from 77 per cent in 1963 to 87 per cent.

Table 1

N.L.L. Unit Costs in £s per issue

	1963/4	1964/5	1965/6	1966/7	1967/8	1968/9
Number of issues	187,593	257,292	336,157	435,750	536,132	651,781
Staff costs/issue	0.53	0.46	0.40	0.37	0.35	0.342
Literature binding costs/issue	0.715	0.52	0.51	0.481	0.493	0.480
Total direct net expenditure per issue (excluding cost of translating programme)	1.31	0.99	0.92	0.867	0.855	0.853

The figures quoted only relate to the expenditure under the N.L.L. "vote". They do not relate to such things as building costs and maintenance of buildings.

This table is prima facie evidence that the N.L.L. is not being unduly pedantic in collecting literature, making records or handling requests. Had the unit costs risen it would not have proved the contrary, but it would have justified a more detailed analysis of the situation.

Whilst I agree that "unit costs" have no precise meanings as regards operational efficiencies it is a simple and useful concept from a management point of view. The unit cost concept will probably be useful in other libraries or in analysing a library complex, and it is interesting to note that the Dainton Committee made use of this concept in analysing the national libraries problem. Of course, overall unit costs given in Table 1 could be broken up in different ways (e.g. on recording or on issuing) but this has only been done at the N.L.L. for special investigations.

In introducing this concept of unit costs I find that I have already introduced another concept, that of the success factor. This to my mind is a very simple concept which is fundamental to the operation of a library but, surprisingly enough, it is rarely used. At the N.L.L. the figure is obtained by dividing the number of issues by the number of requests received excluding only those which have been returned to the user for further information or action. The success factor by itself is a crude guide for comparing the services of different libraries or in different parts of the same library. I appreciate that in some libraries, particularly those with open access, it is a little more difficult to measure than it is at the N.L.L. Nevertheless, it seems to me so fundamental that it should be measured more often.

Of course, by itself the success factor does not indicate what actions to take but it does indicate to some extent whether previous actions have been successful. To obtain further information it is necessary to analyse the extent to which the requests are satisfied in more detail. At the N.L.L., for instance, we found that for English language books the success factor was rather low for relatively recent titles. Part of the trouble we found was due to requests arriving at the N.L.L. either before publication had actually taken place, or very shortly after publication. There was obviously little we could do about the former, but we have taken action to change our acquisition procedures to obtain some books from publishers on publication by placing standing blanket orders for them.

The "success factor" is a measure of the ability of the stock to meet the needs of users eventually. It does not indicate how long users may have to wait. The main cause of waiting at the N.L.L. is due to items being on loan when required. To obtain a measure of this trouble we measure the percentage of issues which are made via a waiting list procedure. The policy of the Library has been to duplicate holdings to try and keep the percentage on loan when required down to not more than 10 per cent over-all. This level was decided somewhat intuitively after considering the experience of the Science Museum Library. It appears to be acceptable to the N.L.L.'s users.

The problem of keeping down the waiting list is mainly, but not entirely, one of duplicating current periodicals and books. In deciding what to duplicate the basic assumption is made that the demand is random in the mathematical sense. This assumption leads to a mathematical model in which Poisson's distribution plays a part. Using this model it is possible to examine the effect of duplication using different criteria. This model can thus be used for developing a formula for relating the number of copies of a title ordered or retained to the demand. The assumption of randomness of demand appears to fit most of the variations of demand which have been observed for recent volumes of particular titles. Fortunately, if the assumption is incorrect and a demand is orientated (i.e. it sometimes concentrates on one item and sometimes on another) then a greater degree of duplication would be required to attain a good level of service than Poisson's distribution would suggest. So that the mathematical model satisfies the administrative requirement that duplication should be on a conservative basis.

In applying these ideas, particularly to periodicals, I have developed a simplified approach which has reduced the problem of deciding how many copies are required to one of mental arithmetic. Since more elaborate treatments have been set down it might be useful to outline the method used.

First the concept of "availability of demand" "A" is defined as the percentage of time one or more copies are available on the shelf. For a title held only in a single copy "A" is the fraction of time the publication is on the shelf. "A" is measured having regard to the number of loan issues. It thus measures the availability for other purposes (e.g. for the photocopying service).

For multiple copies the demand is thought of as divided equally between the copies. Thus, if each copy would have a shelf availability of "a for n copies the non-availability (i.e. $1 - A$) would be given by $1 - A = (1 - a)^n$.

The application of this idea involves the concept of "the average demand". This could, in fact, only be estimated after measurements over a long period, and these would be useless because the total demand on the N.L.L. and the demand for any particular volume of a title are each changing. It is thus necessary to make the observations over a short period and treat the observations conservatively having regard to Poisson's distribution. Thus, if a demand of 15 was observed in a particular period this might mean (allowing for two standard deviations) that the average was 9 to 25. The lower figure would be used to determine whether to duplicate or not.

In deciding whether to try and use the secondhand market to fill gaps in the holding of periodicals the number of requests for the volumes not held have been the major consideration. The justification of this activity is quite simply that experience has shown:

a) the older volumes which have been more frequently asked for, are more likely to be asked for in the future,

and b) those older volumes for which there is some demand are more likely to be available from the secondhand dealers.

(N.B. the majority of the N.L.L.'s secondhand orders are below the demand level which produces reprint editions.)

Where use data has not been available - as for instance for social science serials - the N.L.L. has used citation data to decide which back runs of serials to seek from the secondhand market.

The figures I have referred to have related to the library's main operations. The N.L.L. has made a number of other measurements to guide policies. These have related to the need for translations, the use and need for bibliographies, and the methods of operation of the N.L.L.'s loan and photocopying services. These surveys have related to particular practical possibilities. Special care has been taken to avoid asking potential customers what they think they want. I do not believe that users necessarily know what they need or what they would do in some hypothetical situation.

This brings me to a problem which so far I have failed to solve. The problem of estimating the future demand on the library. Nowadays by extrapolation on the "demand/year" curve we can obtain a reasonable guess of the likely demand this year or next. This process however leaves me with an uneasy feeling that I do not know what is really happening. Extrapolation may be very misleading if I try to estimate the demand in five or ten year's time. The main bulk of the N.L.L.'s ultimate users - the U.K. scientific manpower - is increasing at about 5 per cent per year. If the increasing demand on the N.L.L. was about 5 per cent per year it would be enough to assume that the future demand would follow the growth of scientific manpower. But the growth in demand is about 20 per cent. Some other factors are clearly at work. It is probable that the library habits of the organisations we lend to are changing. I do not think market research would help very much. Some new factor such as devaluation may in time have a major influence upon the situation. So I am driven to guessing the future demand. This appears to be a problem which faces new services, for new services create new demand.

The ideas I have set out so far are rather simple yet it seems to me that they are sufficient to form a basis for the planning of a national libraries service. Indeed the basic proposals of the Dainton Committee depend upon something even simpler - the fact that the cost of accommodation in Central London is about five times that in Boston Spa. If all the activities of the N.C.L. were transferred to accommodation costing the same as at Boston Spa using the figures given in the Dainton Report (mainly in Appendix 4),(3) the position would be as given in the following table.

Table 2

Unit Costs 1967/8

	N.L.L. Issues from stock	N.C.L. Issues from stock	N.C.L. Interlibrary requests
No. of requests satisfied	536,000	21,000	78,000
Success factor			
Unit Cost per request satisfied in £s			
Staff costs	0.35	0.88	0.99(1)
Literature etc.	0.50	0.54	0.04(1)
Accommodation at Boston Spa prices	0.30	2.34	0.34(1)
Total	1.15	3.76	1.37(1)

(1)N.C.L. costs only

The bases on which these figures have been compiled may be slightly different. The N.L.L. staff costs include some indirect costs (e.g. the cost of staff required to provide indexing data for MEDLARS), and the N.L.L. figures exclude the receipts from photocopies although they include the cost of photocopy staff and postage of photocopies. However, the figures given in Table 2 are sufficiently comparable to permit some general observations.

To begin with the N.C.L. interlibrary issue cost has to be increased to allow for the costs in the library which actually issues the required item. In addition, to obtain some comparable figures, it is necessary to include the costs of co-operating libraries in providing data to the N.C.L. and of dealing with requests for items not available. It would appear most improbable that the external staff costs per N.C.L. request satisfied by another library would be less than the N.C.L.'s own costs. If the two costs were equal we would have, using the Boston Spa accommodation costs:

Unit cost per issue (1967/8)

From N.L.L. stock	£1.15
From N.C.L. stock	£3.76
N.C.L. interlibrary request	£2.25

In the existing situation and with the present procedure it would appear that in general it is appreciably cheaper for the N.L.L. to continue to satisfy the requests it receives from its own stock using the existing acquisition policies than to rely on interlibrary requests.

There is another deduction which can be made about the N.L.L. from the figures in Table 1. Allowing for the depreciation of money the unit costs are falling. This is the situation which should exist in an efficient industrial organisation with rising output. It does not happen, however, if a unit has passed the optimum size for its existing methods of operation. Thus, the evidence suggests that if there is such a thing as an optimum size for the N.L.L.'s sort of activities the N.L.L. has not yet reached it. This observation I find re-assuring, for - as far as I am aware - the N.L.L. is already dealing with more requests from remote borrowers than any other library in the world and the volume of demand is still rising rapidly.

The data available in the N.L.L. makes it possible to examine its activities in more detail using unit cost ideas. For instance, it appears that the present policy of the N.L.L. of purchasing back sets of periodicals for which there is some measurable demand is still justified on a unit cost basis. Also, it appears that the present policy of the N.L.L. of accepting a majority of back sets of periodicals offered to the library is justified on a unit cost basis. Using the Dainton report assumptions the nominal accommodation costs per issue of the material in the N.L.L.'s pre-1960 serial store is about £0.65. Moreover, this unit cost is falling as the number of issues from this store is increasing.

The position regarding the N.C.L.'s activities is not so easy to interpret. Even if the N.C.L. were operating in accommodation which cost no more than that at Boston Spa, in the present situation it would on average be cheaper to use the Union Catalogue than to build up a central collection. Of course, the existence of a central loan collection would alter the demand. Unfortunately, as I have mentioned earlier, the problem of estimating the future demand on a new service has not yet been solved satisfactorily. A central loan collection in the humanities would attract a great deal of demand which now goes to the regional library bureaux. In addition, the central loan collection would create some new demand. There appears to be, however, insufficient data at present available to estimate what the total increase in demand would be. Centralisation of the existing arrangements could produce a more efficient system from the user's point of view, but how the costs of a centralised system would compare with those of the existing system is not clear. Whatever happens it is probably going to be difficult to establish on economic grounds the need for a separate extensive central loan collection of foreign language books. Of course, such a collection might be justified from a service point of view.

This problem illustrates the basic dilemma we face in planning library services. Whilst the unit cost concept is clearly useful, it does not permit any allowance to be made for the level of service. This dilemma will exist until there is some method of measuring the cost of not having an item when it is wanted.

This paper has concerned itself with library planning using existing technologies. The book as we know it has survived for over 2,000 years because it is a convenient format. It is true that there are those who have forecast its early replacement by the format produced by some new technology. However, I suspect that the book as we know it will survive a little longer. If the book is to be replaced, this will only come about if the new format is technologically reliable, easy to use and, above all, economically justified. The situation could change quickly, but the replacement of the book by some new format is, at present, no nearer than the horizon. However, it is not impossible that, before long, it will be possible to have rapid access to books stored in remote places. The decision whether such a system should be introduced is likely to be mainly determined by unit costs.

N.L.L.
June, 1969

D. J. URQUHART

1. URQUHART, D.J. J.of Doc 13 pp 13-31, 1957

2. Royal Society Scientific Information Conference, Report. p.103, 1948

3. Report of the National Libraries Committee, Cmnd.4028, H.M.S.O., 1969

Discussion

When asked whether the reason why the level of external borrowing was rising proportionally faster than the user population was that the same users were reading more, or that the reading level of previously inactive users was being raised, Mr. Dammers replied that he did not know; he could only say that current awareness services had been improved and consequently readers were better informed of relevant papers. Mr. Mackenzie observed that Mr. Dammers' readers were using about 18, 000 papers per year, and in 5 years this would have increased to about 40, 000; he asked if they would have enough time to do all this reading. Mr. Dammers agreed that the readers would collect more papers but would not necessarily read them, and he pointed out that a great deal could be assimilated from a paper just by glancing at it. Also, the present number of papers being read was not very high - 10 papers per week per reader. Mr. Woodward said that in the last few years the annual rate of growth of inter-library loans at Durham was 30%, and he thought that part of this increase arose because readers who previously did not read much had now begun to catch up on their reading. Dr. Urquhart suggested that when more abstracts and selection devices were acquired users would make an initial selection better suited to their needs, and this in fact represented a wider scatter of the literature; consequently the library was less able to meet their requirements from its own holdings. Mr. Fairthorne said that when a library improved its facilities its readers would make use of them and ask for more literature, but this did not necessarily imply that they would be doing more reading; often they would merely use the facilities to find clues on what to read.

Mr. Dammers said that in spite of what had often been said elsewhere about the direct distribution of papers to individuals he could not see the disappearance of the journal, because it was such a convenient way of bringing together papers on related subjects; but he could not foresee where the journals would be kept. Perhaps special libraries would collect the small number of important journals in a particular subject field, and these would be backed up by large central collections.

Referring to the size of the average special library in terms of number of periodicals held, Dr. Urquhart said that he thought that similar data was given at the 1948 Royal Society Scientific Information Conference, and that the average size had not materially changed since then. As special libraries grew, new ones were formed and the average size remained constant. Mr. Vickery said that data at Aslib suggested this to be true. Mr. Dammers was asked if the figures he gave for the number of periodical titles held in relation to the number of volumes had a more general application and if they could be applied to university libraries; he replied that he did not know and that he only used the figures for guidance in the development of his library. Mr. Fairthorne said that some work he did five years ago, relating the number of periodical titles held to volumes, agreed with the figures quoted for special libraries, and that deviations from the curve really were exceptional libraries, such as the library of a motion picture company.

Professor Leimkuhler was asked if he had any data on the minimum average cost which results from a library of a standard size in a certain situation, but said it was not possible to answer this. Mr. Vickery said that in practice the concept of minimum average cost had been carried out more in special libraries than in university libraries because the former had discarded more. Dr. Urquhart said that from his experience he thought that university libraries had not yet heard of unit costs. Mr. Mackenzie agreed, but said that one of the difficulties of costing in university libraries was that it was not possible to define precisely the aims of a university library, whereas the average special library had as a target the overall profitability of the firm. Cost benefit analysis of a university library was not going to lead very far until its educational and archival purposes had been defined.

Dr. Urquhart said that requests for periodicals from a store of pre-1959 material at NLL were increasing, but there was no exponential decay in the usage of books in that store. Mr. Buckland commented on this apparent lack of exponential decay, and suggested that this pattern might nevertheless exist in terms of the proportionate distribution of demand, but that the steep rise with time in the level of demand would mask the exponential curve in terms of the absolute number of requests for individual items.

Mr. Fulford asked Dr. Urquhart if he agreed that providing a loan service in the sciences was a different problem to providing a loan service in the humanities, and if it was feasible to compare the unit costs. Dr. Urquhart replied that if the requests which at present were satisfied regionally were satisfied by NCL, the use of NCL stock would be much greater, and the corresponding unit costs much lower; but there was no fundamental difference between science books and arts books, and providing a loan service for them should be a similar problem.

Mr. Longworth wondered what was going to happen to publishers as more and more information was stored in computers in digital form. Mr. Fairthorne said that it would not be economic to store large amount of printed information in digital form in the foreseeable future. Mr. Dammers disagreed, and said that mass storage of literature was not a long way off since store sizes were increasing all the time; it would happen in science first because the economic pay-off was quite obvious. He was asked for more details of machines and their sizes, and he mentioned the IBM photo-digital store with a capacity of 10^{12} bits which would accommodate 100, 000 volumes. Mr. Duchesne observed that this was not in fact very large when one though university libraries. Mr. Mackenzie asked how computer storage would affect a university of 5, 000 students: would they all have immediate access to the store and have their textbooks printed out? Mr. Dammers said that the process would obviously be gradual: scientific literature might well be stored first, and the university library would have access to national stores; subsequently departments would have access to a store in the university library. He added that it was bound to happen eventually, and university libraries had to face up to the problem now and not turn their backs on

In reply to the question of why he foresaw the increase of inter-library loans in the future rather than a greater increase in the purchase of literature, Mr. Dammers said that this happened to be more economic; the tendency appeared to be for special libraries to hold the basic journals and rely on inter-library loans for the others. Mr Dammers was asked where he got his inter-library loans from; he replied that they came mainly from NLL and the Chemical Society, and added that at the moment NLL could cope very well, but he wondered how it would fare in the early 70's. Dr. Urquhart replied that even with present techniques NLL could cope with an increase in demand by a factor of 10.

Mr. Snape asked Professor Leimkuhler if the factors affecting the demand for literature of the humanities were different from those affecting the demand in the sciences. Professor Leimkuhler agreed that there was a difference, but he had not been able to express it in an analytic way. Dr. Urquhart said that one of the differences was the time scale in the fall-off of use: it was very rapid for the sciences and less so for the humanities. This meant that decisions on duplication of books were more difficult for scientific literature, because in the sciences the fall-off in demand was so high that by the time the need for duplication had been recognised and the duplicate bought the demand had passed; therefore the demand had to be predicted, and this was done at NLL by regression analysis. In the humanities, with a less rapid fall-off, it was possible to wait for a demand to appear before buying a duplicate. The need for duplication of periodicals was easier to recognize, and periodical literature was more important in the sciences than in the humanities, this was another factor which made provision easier for the sciences. Mr. Fulford said that half of the requests at the British Museum were for books published before 1919; Mr. Buckland suggested that the British Museum was a special case because it catered largely for a residual demand.

Mr. Brookes asked Mr. Dammers what the unit cost of providing searches on 150 profiles was. He replied that it was around £35 per profile per year, run on C.A. Condensates tapes. Mr. Brookes asked Mr. Mackenzie if he thought it was feasible to provide such searches in university libraries at that price; he replied that he thought that it was. Dr. Urquhart said that a MEDLARS search costs about £30 per year for a profile search, but the system was not primarily designed for profile searching. Dr. Hawgood observed that one tape was suitable for all Mr. Dammer's users; unless there was a tape covering all subjects of interest to the university users, he could not see how universities could provide a similar service. Mr. May said that OSTI were encouraging the varuous bodies (INSPEC, UKAEA, etc.) who were working on SDI programmes to agree on a common system, so that a programme for searching an INSPEC tape could also be used for searching others, e.g. NSA users. Mr. May asked Mr. Dammers if his profiles were individual or group profiles, because NSA had found that a cost reduction by a factor of 10 could be achieved by changing to group profiles and selecting them from a standard list. Mr. Dammers said that this might happen in the future, but at the moment he wanted to put as few barriers as possible between the user and the system, and therefore all his profiles were individual profiles.

Mr. Mackenzie asked how was it possible to achieve a balance between acquiring literature and providing current awareness services: would the humanities suffer if the university library supplied current awareness services to scientists? Mr. Brookes said the scientists in the University of London, who had previously been accustomed to such services in industry, had set up their own services in the university, and often these duplicated each other and were very costly; consequently university libraries could not afford to ignore developments in SDI services. Dr. Hawgood said that universities might soon start charging for computer services and wondered if anyone had investigated the possibility of charging for library services. Replying to the suggestion that, when allocating their resources, university libraries should take into account their readers' views on what services they expected from the library, Mr. Morely said that studies in economics had shown that it was not possible to allocate resources within a community on the basis of its members' preferences. Dr. Urquhart said that the NLL had tried to assess what would happen if charging for MEDLARS searches was introduced: it was estimated that the demand would fall by 55% and that most of this would be due to a decrease in demand from universities.

Mr. Buckland asked Mr. Longworth what effect the implementation of the recommendations of the Maud Committee's Report would have on public libraries; he replied that, if the recommendations were implemented in their present form, they would lead to an enormous upheaval in the structure of public library services. Lancashire would lose 20% of its services to Manchester, 20% to Liverpool, and the remaining 60% to boroughs in North Lancashire.

Mr. Longworth asked Professor Leimkuhler how was it possible to assess the benefit of a particular loan; he replied that this was possible only in terms of the cost to the user of obtaining that particular loan. Dr. Hawgood said that there were two ways of doing this: one was to calculate the macroscopic cost and the alternative open to the library manager, and the other way was to examine the various alternatives open to the borrower and why he chose to spend his time in a particular way. Dr. Urquhart said that public libraries should collect statistics of the proportion of issues which were educational, recreational, research, etc., because some people thought that public libraries were entirely recreational; the publication of such statistics would help to define the rôle which the public libraries were playing in the community, and in particular it would emphasize their educational function. Mr. Longworth agreed that this was one area in which there was remarkably little information.

SESSION 2: TECHNIQUES OF ANALYSIS

C. F. Carter	The allocation of resources in higher education
R. M. Morley	Maximising the benefit from library resources
Discussion	

The allocation of resources in higher education

CHARLES F. CARTER

I am told that at the headquarters of the Department of Education and Science in Curzon Street there is a doctrine called 'Carterism', and even a variant called 'modified Carterism'. My academic colleagues will, I fear, assume that this doctrine is concerned with making true academic values subservient to the arid calculations of the economist. But this is not the point at all. We are short of resources for higher education. There are times when the only right thing to do, if you are short of resources, is to demand more. The present, however, is not a moment at which such demands stand much chance of success: to prove that one needs to be more expensive, when the country is notable for its economic failures and the Secretary of State has at hand a possibly cheaper alternative in the other part of the binary system, would not be likely to yield much satisfaction. Consequently we must look to the best use of the resources we already command. This is not an attack on academic values: it is their defence in times of hardship.

For, of course, the making of choices about scarce resources which have alternative uses is inescapable. Although many of you are, I am sure, dedicated to the belief that a library can never be too large, it is obvious that a university which spent all its money on its library and nothing on anything else would be somewhat ineffective. The answer to the question 'What proportion of recurrent income should be spent on a library?' is therefore not 100%. But is it 20%, or 10%, or 6%? Is there any rational ground for choosing between these?

In order to answer such questions, we have to have a 'model' of a university*: that is, to define what it is intended to produce, to set out the technology (or the alternative technologies) used to produce the output, and the resources available as inputs. The difficulty about doing this is that the model is insufficiently defined. A university produces -

(a) Services of education and training to first degree students

(b) Services of education and training to higher degree students, attenders at short courses etc.

(c) New knowledge and new developments of culture or civilisation.

(d) The service of preserving and rediscovering existing knowledge and culture.

continued ...

---

* Or, of course, a polytechnic or other institution. The allocation of resources between different types of institution I do not here discuss, though it could be treated in a similar manner.

The library has a special importance in relation to the last function, which is often forgotten because of our habit of making a dichotomy of 'teaching' and 'research'.

Within the two teaching functions, the balance between different types of course is seldom dictated from outside the university. It is loosely constrained by the preferences of entrants, which in turn may be loosely related to ideas, commonly formed on partial, misleading or out-of-date facts, about the ultimate value (financial or cultural) of the education they choose. But within wide limits what universities do to determine the balance of courses is a sort of academic guess, unrelated to market forces. The guess emerges from whatever the decision-making process in the university happens to be. Even so, however, it may be possible to reallocate resources so as to give a new solution (within the constraints) which is more acceptable to the decision-makers. This is a long way off saying that it is 'right' in any absolute sense, but it is a move towards a greater satisfaction of a dominant group.

Similarly, the relation between the 'teaching' and 'research' functions is only loosely determined. It is commonly supposed, though without any very strong evidence, that if research falls too much the quality of teaching will decline: because this is widely believed, and because research is to some people more fun than teaching, a single university which allowed the proportion of research to decline might well suffer by an inability to attract good staff. At the other extreme, an institution which has too much research runs a danger of being regarded as so expensive that it will fail to attract adequate grants. But within a wide range, the research proportion is whatever the university decides, or more often, allows to happen without a conscious decision. Since most university research is a long way, conceptually or in time, from a relation to the judgements of the market place, the allocation of resources between different types of research is almost totally arbitrary or accidental. Pathetic attempts are sometimes made to develop a national 'research policy', sometimes based on the illusion that a nation can and should do a little of everything; but, seen at the level of a particular institution, the balance is decided by the numbers and powers of advocacy of those who influence decisions.

The preservation function is even less well determined, since its very existence is hardly appreciated. It is true that the margins of indeterminacy in all these product areas can be somewhat reduced by looking at the university system as a whole: for instance, this introduces the additional constraint (not felt by any one institution) that the system must produce enough good graduates to provide its own future staff. But the attempts

continued ...

to lessen the indeterminacy by manpower forecasting have had little effect. Future manpower needs can only be very imperfectly described in terms of the demand for particular kinds of specialised training, if only because there are so many jobs for which prior training matters much less than character, flexibility and willingness to learn.

In the model of a university, therefore, there will be large elements which must be labelled 'historical accident', or 'decided by professorial prejudice', or 'vice-cancellarial whim'. But this does not mean that improvement is impossible: simply that we cannot identify an absolute optimum. It should first be our aim to free the model from contradictions (that is to say, features which now or later will cause the violation of a constraint), and then to consider various marginal shifts of resources and set up principles for deciding whether a shift is or is not an improvement. These principles may not be capable of absolute justification: we may have to accept a crude test, such as a majority view of those affected by a change (assuming all opinions to be equally valuable). However, many affairs must be run in this way, in the absence of more refined methods.

The resources to be allocated can be represented by sums of money, to be subtracted from this purpose and given to that: or, since the inputs into the model themselves have alternative uses, we may want to consider a shift of 'hours of staff time', or 'hours of computer time', or 'hours of building use' from one purpose to another. (In the case of buildings, the shift from not being used at all is particularly important.) There can be no rational discourse on these matters without a prior study of the facts of existing use. This is why the Committee of Vice-Chancellors has set up studies of the structure of departmental and university costs, the uses made of staff time, and the uses of buildings and of expensive items of equipment. Unfortunately it is difficult to get comparable data from the public-sector institutions.

It takes time to obtain facts, and those we have so far provide no more than preliminary suggestions. Furthermore, the study of present practice must take account of Parkinson's laws: if you give a department too much staff, it will develop activities to justify the staff it is given. It will not be easy to prove that these activities are a waste of time. The misallocations which are easily observed, therefore, are less than those which actually exist. The kind of information coming forward is as follows:

(a) In some subjects there are significant economies of scale in undergraduate teaching, over a fair range. The probable meaning of this is that, where there is a strong professional expectation about the range of topics to be taught, departments with only a few students have to be relatively over-staffed, and it looks as though a rough definition

continued ...

of 'too small' is 'having less than 120 equivalent full-time students[+]'.

By this standard quite a number of British university departments are too small: if, through lack of demand, they cannot be expanded sufficiently in the course of general growth, and if they cannot plead the justification of providing an essential ancillary service to other departments or of maintaining an essential cultural element within the university, it is probable that the money spent on them, if used elsewhere in the university, would yield more satisfaction. Of course, all such departments will plead justification: the real lesson is for university planners - don't start a new activity unless you are confident that it can reach a viable level.

(b) Some universities have allowed very great disparities to occur in their treatment of different subjects. This might be justified if selective heavy spending was creating 'centres of excellence', but it is not easy to trace a clear relation between heavy spending and quality: most of the instances are probably historical accidents which have been allowed to persist because data systems have not revealed their existence.

(c) On the other hand, some places allocate staff, technicians, secretaries etc. with great apparent fairness, in relation to the student load. But such pro rata allocations are only appropriate if all departments are the same size, or there are no economies of scale: otherwise they imply that all benefits from larger departmental size rest in the department, instead of being shared with the rest of the university. A more abstruse point is that such calculations commonly use 'weighted' student numbers, giving a greater weight to graduate students, and there is increasing evidence that the weights used are much too high - thus giving a bonus to departments with a lot of graduate work at the expense of the rest.

(d) We shall soon know a lot more about the way in which university academic staff use their time. This will not give an exact allocation between teaching and research, because some time is used to the benefit of both these activities: but at least we shall be able to see roughly how the division of time varies between different universities, faculties or grades of staff, and this will give a firm basis for arguing about what it ought to be. At present time is often assumed to be divided fifty-fifty, this being the plausible assumption to make if one is totally ignorant about the truth.

continued ...

[+] i.e. a student spending a third of his time on a subject counts as a third.

(e) We can now document in considerable detail the use made of various kinds of university building. It is plain that in a great many cases this use could be substantially increased, even by measures which do not involve any reorganisation of the university year. However, it must be remembered that a university uses labour much more than capital: if capital costs were properly charged and amortised, they would not be more than 10-15% of total costs. Furthermore, universities have hitherto had absolutely no incentive to economise on capital, because any savings made have not been transferrable to other areas of expenditure.

(f) There is of course a national advantage in economising on capital, which is especially attractive to the Treasury because one might be able to have a period in which very little was spent on university building. Studies have been made, both in this university and nationally, of the possibility of achieving a big saving by rearranging the university year for continuous operation. These studies reveal technical difficulties which are much more severe for a three-year degree than for an American four-year degree: and the present evidence is that this particular reallocation of resources would not be worthwhile.

Let me, however, return to the problem which I mentioned earlier: how much should be spent on a university library? The element here which is most readily assessed is, I think, the requirements of students on set courses. Even this, of course, depends on how many books they buy: it would be logical (though, I fear, impossibly unpopular) if, instead of an indicated but seldom used sum within the maintenance grant, standard textbooks were available on prolonged loan from the library, which would thus be able to assess accurately the number of copies needed. No system will, of course, be proof against the folly of a lecturer who tells a class of 300 to read an obscure article, existing only in a single bound volume of a journal: but generally, even allowing for our lack of knowledge of what students buy for themselves steps can be taken to regulate book purchase and lending policy so as to maintain a prescribed level of consumer satisfaction, which can be measured in terms of success in obtaining a book on demand or within a short waiting period.

Generally, however, university libraries are much larger than can be justified by immediate student needs. What can we say of the function of preserving knowledge and being a repository of culture? It might

continued ...

seem that all that is needed is to build a sufficiently large vault or pyramid in Bloomsbury or Boston Spa, within whose burial chambers would be interred a collection more splendid than we have yet seen. However, culture is really preserved, not by books alone, but by the interaction between them and generations of scholars. This means that the 'preservation' function can be treated as closely similar to the provision of reference material for research. The question to be answered is - how much should be provided locally for a particular community of scholars, and how much can be distant by a few hours in some central place, or available centrally within a day or two in loan or copy form?

I do not think that we can give sensible answers to these questions without more accurate knowledge of how a scholar in a particular discipline uses a library. Scientists commonly think of this use in terms of 'information retrieval': that is to say, a precise question is put to the library (or, in future, to the computer), which presents an answer. To what compound does this spectrum relate? Has anyone synthesised this substance before? It looks as though a great deal of this is capable of centralisation, without serious waste of time. In addition, of course, the scientist wants to 'leaf through' abstracts of recent work, so as to see what is going on and what significant discoveries have been made: this argues for a local library having a full complement of abstracting journals, and a quick access to loans of particular journal articles. But a student of literature or of history is likely to want to refer, in a short time, to a large number of works: different editions of a particular book, related material about the period, primary and secondary sources which exhibit the meaning of some obscurity. This process is intolerably slowed down if each peripheral enquiry involves a reference or a visit to another place; and the question being asked may be much more difficult to put in a precise form than in the case of the scientist, and therefore computerisation may yield an unmanageable mess of references. Technically a collection of printed books, convenient for 'browsing', is hard to beat as material for the kind of research which involves numerous ill-defined peripheral enquiries. It would still be possible for the scholar to concentrate his teaching in one part of the year, and then migrate to an attic in Bloomsbury (or travel daily to some great provincial library) so as to have access to a wealth of material for research. Some will do this: others will be prevented by their ties of responsibility: others will find it intolerable that an idea cannot be brought to the anvil, white-hot, to be fashioned into a paper, but must be put away until the next time there is a chance to travel to a good library.

One could found on such arguments a case for a substantial library in the place where scholars (of certain disciplines) live and work: but a lot more information is needed about the habits of scholars before one

continued ...

could translate this into a system for quantifying how large the library should be. The worthwhile world literature in economics probably costs about £5000 per year, excluding the 130 principal journals and a very large number of Government publications: if a library bought all this, one could discover over a period which publications were never used, and which were used so rarely that the cost of shelf-space and management would exceed the cost of sending users by Rolls-Royce to London and boarding them at the Hilton. The difficulty, however, is to infer what might have been wothwhile among things which have not been bought: and even a guess at this requires a close study of the actual research habits of economists.

My general conclusion, therefore, is that it is a worthwhile exercise, in defence of the true values of higher education, to attempt a more systematic allocation of resources: but that our efforts to do this must for the present be of a very preliminary, partial and tentative kind, partly because the theoretical problems of allocation are difficult, but mainly because the substructure of facts is incomplete, and can only be built slowly. Still more difficult is the allocation of national resources between different types of higher education, which at present must be largely a matter of guesswork or political preconception. I hope, therefore, that those who receive tiresome enquiries into the facts of their academic life will appreciate that in the long run they will be helping to defend and improve the academies they serve.

Maximising the Benefit from Library Resources

by R. Morley

University of Durham

This is a report of part of the work at the University of Durham Computer Unit of the Project for Evaluating the Benefits from University Libraries. The Principal Investigators of the project are Dr. John Hawgood, Director of the Computer Unit, and R. Morley, Lecturer in Economics. The team includes Mrs. Jean Hopkins, M.G. Ford, and Lt. Col. W.E.M. Morris. Credit is due to the team; errors in this paper are due to the author.

We are grateful to the Office of Scientific and Technical Information for their financial support and their encouragement to make ourselves intelligible, to the librarians at Durham, particularly Miss A.M. McAulay, for their invaluable advice and assistance, to our fellow researchers at the University of Lancaster and elsewhere, and to the staff and students of the Universities of Durham and Newcastle for their participation and forbearance.

Maximising the Benefits from Library Resources

The usual approach to information problems is to find out the needs of students/teachers/researchers and then attempt to provide for these needs as cheaply as possible. This has the merit of treating the present members of an organisation as the most important. It takes into account the fact that the reproduction of research findings is far cheaper than their production. In practice it does not work very well, partly because the needs are difficult to discover, but also for other reasons.

Most organisations, and universities in particular, are severely limited in their actions by decisions taken in the past. A university library has a building, a bookstock and a team of skilled workers. None of these resources can be changed significantly on a month-to-month basis. Of course, small re-arrangements to the building are possible; proportionately small additions to the collection of monographs and serials occur continuously; there may even be some change in staff. However, an expansion of students/teachers/researchers which necessitates a new library is a major top-level decision within the university. So is a change of fields of study large enough to render the existing bookstock useless. All this suggests that an alternative approach may be worthwhile. Instead of taking needs as given and finding how to meet these needs at minimum cost, we can take resources as given and find out how to maximise the benefit from these resources.

The "needs" approach requires a continuous re-arrangement of the fruits of the past to ease the research of the present. The "maximising" approach requires some adaptability by present researchers in order to gain from the fruits of the past. Perhaps the latter is equally realistic. If one is presented with an island of seagulls' droppings one can hold one's nose or one can find the nitrogen content, but the clever man will do both.

This paper is about the "maximising" approach. It does not pretend that this is superior to the "needs" approach, but rather a complement to it. The "needs" approach is essential for long term planning; the "maximising" approach seems preferable for month-to-month planning and it yields some insights into the preferences of library users which may contribute to an understanding of their needs.

In Part I a library in fairyland is discussed in some detail. This unrealistic but simple example allows many of the principles of linear programming to be outlined, and some illustrations of the usefulness of the technique are given. The approach is seen to rely upon a knowledge of the relative importance of the different library activities and two methods of measuring this are proposed in Part II.

In Part III the method is used on a real example, the Arts/Social Science Library in the University of Durham. The implications of changes in users' preferences, or changes in the resources available, or changes in library technology can be assessed by a few seconds' time on a computer.

I A library in fairyland

This library engages in two activities: answering queries and lending books. There is no shortage of books and no shortage of space. The only resources that are scarce are three different sorts of labour: senior librarians, junior librarians and clerical labour. The problem facing the library is to decide how much of these two services it should provide each day, given that resources are limited. The library is motivated to find the best combination because it receives 7 shillings per query answered and 10 shillings per loan, and it wants to maximise profits. (The prices are an indicator of the strength of demand for the services. Where they came from is another matter! We can assume for the moment that there are many libraries competing with each other, but in Part II this assumption will turn out to be unnecessary.)

First the maximum amounts of queries and loans which the library could produce are found. This involves finding how much of each resource is available and how much of each resource is needed to produce one unit of each activity. There are 980 minutes of seniors' time available each day, and it takes 11 minutes of this to contribute to the answer to one query and 2 minutes to contribute to the provision of one loan. The constraint on library activities imposed by the shortage of senior librarians can be described as

$$11\,Q + 2\cdot L \leq 980$$

when the numbers refer to minutes of seniors' time, Q is the number of queries answered in a day, and L is the number of loans made. The total commitment must be less than or equal to the total available. Both junior librarians and clerks are also needed before the activities are completed, but the constraint imposed by only the shortage of seniors is illustrated in Figure 1. Any point outside the triangle is impossible because of the way the problem is defined. Any point inside is inefficient because with the relationships as given more could be produced. Only on the borders of the triangle is production both possible and technically efficient.

There are 400 minutes of juniors' time available each day; it takes 3 minutes of this time to contribute to answering a query and 5 minutes for a loan. There are 100 minutes of clerks'time available each day; it takes one minute of this to contribute to a query and one minute for a loan. All this can be summarised in Table 1.

Table 1

Resource	Amount needed per unit of Q	L	Amount available
Senior	11	2	980
Junior	3	5	400
Clerical	1	1	100

In Figure 2, the three constraints are shown. The border oabcd encloses the possible combinations of Q and L. This area, including its boundary, is the set of production possibilities. Points to the north-east of ab are impossible because there is not enough clerical labour, of bc because of insufficient junior labour, of cd because of insufficient senior minutes. Points west of oa and south of od are meaningless because negative quantities of Q and L are impossible.

Since the library receives 7 shillings for each query answered and 10 shillings for each loan, the total profit will be 7Q + 10L. The next table shows the combinations of Q and L represented by the corners of the set of production possibilities (all numbers are rounded down) and the profit which results from this combination.

	Q	L	Profit
a	0	80	800
b	50	50	850
c	86	13	732
d	89	0	623

The economically efficient combination is at b, when the library answers 50 queries and provides 50 loans each day.

A change in demand

Suppose that the community served by the library now contains an increased number of technologists so that the demand for the query answering service increases. The price of queries goes up to 11 shillings and the price of loans stays at 10 shillings. The library is now trying to maximise the profit function 11Q + 10L. When Q = 50 and L = 50 the profit increases from 7Q + 10L = 850 to 11Q + 10L = 1050 shillings. However the ratio of prices has now changed. When this happens the library hunts around for another blend of activities to see if a reduction in L and an increase in Q would be worth while. At c, in Figure 2, 11Q + 10L = 11(86) + 10(13) = 1076, so the new blend is more profitable at the new prices.

In Figures 3 and 4, the two profit functions are illustrated. Each function is a family of lines of equal slope, and the choice of which blend of activities to produce depends on the slope. Note that the slope depends on the ratio of prices rather than on the absolute prices. The price ratio shows how the community would be prepared to swap one activity for the other, and this is all that is needed in order to decide what blend of activities to produce from given resources. The total profit is useful in order to decide whether to expand or contract the library as a whole; it shows whether resources are better used in the library or rather in other organisations in the community. This difference between relatives and absolutes becomes very important when we descend from fairyland.

An increase in resources

An increase in clerical time available from 100 to 112 minutes a day allows the production possibilities to be increased to Q = 83 and L = 29. In Figure 5 the increased level of this resource is shown by the dotted line. At the point e all three types of labour are used fully, as shown by the meeting of the three lines. No type of labour is a bottleneck and none is idle. If trained labour could be hired easily, or if labour could be fired easily without leading to insecurity, this position could be achieved by trial and error. In practice many libraries do seem to be near this point except during a period following some major change.

A more flexible labour force

Both clerical workers and junior librarians require a short training period and a similar background of general education. If there were more flexibility between these two types of labour they could be considered as one type. In this case the two constraints could be merged into one constraint.

$$\left.\begin{array}{r} 3Q + 5L \leqq 400 \\ Q + L \leqq 100 \end{array}\right) \text{ become } 4Q + 6L \leqq 500$$

Figure 6 shows the new single constraint as a dotted line replacing the two constraints, one for juniors and one for clerks. Provided production occurs at b, this increased flexibility does not allow any increase in production. However, as soon as demand conditions change so that a movement away from b is justified, the flexibility leads to a greater output than would be possible when the old division of labour applied.

A change in technology

A new aid to information work allows a reduction in the amount of senior and junior librarians' time required to answer queries. If the innovation were adopted the seniors would need to spend only 9 minutes per query instead of the 11 needed at present; the juniors' time per query would be reduced from 3 minutes to 2. How can the effects of the aid be calculated to find out whether the innovation is profitable enough to justify the costs of adopting it?

The present set of production possibilities was given by:

$11Q + 2L \leqq 980$ minutes of senior librarians' time

$3Q + 5L \leqq 400$ minutes of junior librarians' time

$Q + L \leqq 100$ minutes of clerical time

If the new aid were adopted, the production possibilities would become:

$9Q + 2L \leqq 980$

$2Q + 5L \leqq 400$

$Q + L \leqq 100$

When the prices were 7 shillings for queries and 10 shillings for loans, the mostprofitable combination of activities was shown as point b, but at this point all of the time of the senior librarians was not used up. Therefore any further saving in senior librarians' time would be worthless.

However the saving in juniors' time would allow an increase in production. This increase would be very much more marked if the amount of clerical minutes available were increased at the same time that the innovation was adopted. Removing one bottleneck often creates another which is far cheaper to remove.

The new aid becomes more profitable if the price of queries increases to 11 shillings, so it might be worth while adopting the innovation in anticipation of an increased number of library users who were technologists. The profitability of the innovation depends on which set of prices prevails and how much clerical labour is available.

In a situation where a number of different policy changes are possible, the various outcomes can be calculated and tabulated as in this example; shown in Table 2.

Table 2

Clerical time available	Prices in shillings		Old technique			New technique		
			Optimal blend		Profit shillings	Optimal blend		Profit shillings
	of Q	of L	Q	L		Q	L	
100	7	10	50	50	850	33	66	891
112	7	10	83	29	871	54	58	958
100	11	10	86	13	1076	100	0	1100
112	11	10	83	29	1203	108	4	1228

These examples are given to show the many uses to which a linear programming approach can be put. Unfortunately the approach depends upon a knowledge of the prices which can be obtained for the services of the library. The university is not a market economy and there are no prices. Obviously a university cannot use quite the same methods of planning that were appropriate in fairyland.

II Planning without actual prices

Let us suppose for the moment that we know the range of activities that the library could have produced; we know the actual blend that the library did produce, and we know that the library was economically efficient. Under these circumstances can we deduce the prices that must have been prevailing at the time?

In terms of the example of Figure 2, if we know the production possibilities and if we observe that production occurs at b, can we deduce the prices? At least we know that if b is the economically efficient combination, then the slope of the profit function must be on or between the slope of the constraint ab and the slope of the constraint bc. Now ab is the junior librarian constraint, $3Q + 5L \leqq 400$, whose downward slope is $\frac{3}{5}$; and bc is the clerical constraint, $Q + L \leqq 100$, whose slope is 1. Therefore the profit function has a slope of between $\frac{3}{5}$ and 1. But this slope gives the price ratios. It tells us that if the price of a loan is one unit, then the price of a query answered is between $\frac{3}{5}$ and 1 units.

Alternatively, we might observe that production occurs at c. This is where the clerical constraint bc meets the seniors constraint cd. The slope of the clerical constraint, $Q + L \leqq 100$, is 1, and the slope of the seniors constraint, $11Q + 2L \leqq 980$ is $\frac{11}{2}$. The slope of the profit function must be between 1 and $\frac{11}{2}$. So if the price of a loan is one unit the price of a query answered is between 1 and $\frac{11}{2}$ units.

At least this method sets a limit to the range of price ratios. In some circumstances this range can be quite narrow and can give close approximations to the relative importance of the library activities. If a library is efficient, and if the formal and informal communication systems within a university are functioning properly, then observation of the library allows numbers to be put to the relative value of the various library activities. These numbers are relative rather than absolute, but they provide sufficient information to allow library planning to take place. (More complicated examples can be solved using trigonometric techniques.)

Most people seem to find the idea of swapping easier to understand than the idea of pricing, so discussion is more likely to be informative if it is conducted in terms of swap-rates rather than price-ratios. (A swap-rate is the inverse of a price-ratio. If the ratio of the price of Q to the price of L is $\frac{3}{5} = 0.6$, then users are prepared to swap $\frac{5}{3} = 1.66$ of Q for one of L, or about 17 for 10).

In many cases the range of ratios is so large as to be unhelpful. An alternative approach will be shown to be more precise although it does need more data.

Using economic theory to interpret past decisions

The university library is a major supplier of information services within the university. With any one of the services that the library supplies, the benefit will increase quite rapidly at first as the service passes the threshold of awareness of users and as the users come to feel able to rely on the service and assess it. If the service is expanded further the benefit from it will continue to increase, but less and less quickly. In some cases a service could be expanded to a level where users are over-loaded with information and the total benefit from the service actually decreases. The crucial variable is the benefit obtained from the last unit of the service, the marginal benefit. Although crucial, this variable is not known. However, it can be imputed by examining the way in which costs are incurred in the library (see Figure 7).

Over a period of a year or two the library has the opportunity to adjust the amount of many of the resources which it has available. In the previous section we had assumed that these were fixed because we were dealing with only short periods of time. The total of the resources will depend on the size of the budget, but the mix of resources will change as demand conditions change and call for changes in the blend of activities. Since the library knows the cost of these resources and has a fair idea of the amount of each resource that is needed to produce a unit of each activity, it also knows the cost of a unit of each activity. (In economic jargon, these costs are medium-run average variable costs and we assume here that they are close approximations to short run marginal costs. In cost accountancy terms, they are unit costs which do not include any allocation for overheads.)

An efficient library will adjust the blend of its activities in such a way that £1 worth of resources devoted to one activity could not give better value if it were devoted to another activity. If the library is trying to get the most from a limited budget, it will continue to expand a particular library activity up to the point where a £1 of budget could be better spent on increasing another activity. But this is to say that the librarian (or rather the complex committee system that decides such matters) has an intuitive idea of the benefit to be obtained from the last unit of any activity that is provided. We are assuming that the marginal costs of, and the marginal benefits from, the activities of the library are known. (In Part III a test of these assumptions is given and they pass, so read on!)

Suppose that the marginal cost of activity x is ten times the marginal cost of activity y, and the librarian is allocating variable resources efficiently. In this case the librarian must judge that the last unit of activity x is worth, to the community served, ten times as much as the last unit of activity y. The ratio of marginal costs equals the ratio of marginal benefits. Writing MC for marginal costs, MB for marginal benefits, and subscripts x and y to denote the two activities:

$$\frac{MC_x}{MC_y} = \frac{MB_x}{MB_y}$$

Rearranging and generalising to more than two activities:

$$\frac{MB_x}{MC_x} = \frac{MB_y}{MC_y} \;.......\; = \frac{MB_z}{MC_z} = k.$$

where k is some constant (and on the assumption that libraries are economically useful, k is greater than one).

So $MB_x = k.MC_x$ and the value of the last unit of each activity can be found by multiplying the marginal cost by k.

But if we are planning only small changes in library activities, the marginal benefit can be assumed to stay fairly constant. Writing x, y, z as subscripts to denote which activity, and as variables to denote the level of the activity, we can compare two different blends of outputs by comparing two different numerical values of x, y, z in the following expression:

$$MB_x.x + MB_y.y + ... + MB_z.z$$

which is the same expression as:

$$k.MC_x.x + k.MC_y.y + ... + k.MC_z.z$$

Therefore small changes in the blend of library activities can be assessed by using the expression:

$$MC_x.x + MC_y.y + ... + MC_z.z$$

since this is the profit function as used in Part I. The k can be omitted, because only the ratios are relevant.

The reasoning can be reversed. If the marginal costs are known, the ratio of marginal benefits can be imputed from these.

III Application to the Durham Arts/Social Science Library

Assessing the relative importance of library activities

The choice of library activities to be considered depends on the sort of problems which are relevant to short-term planning. Our choice was influenced by problems such as: What is the relative importance of an inter-library loan compared with purchasing a book? How important are user-services compared with other activities?

The classification of activities must be done in a way that is mutually exclusive and allows quantification. It need not be exhaustive although the following list does cover the major activities at Durham. The list gives those activities which were considered together with the units of measurement. The code letters are provided to aid cross-reference in the tables.

I	Increasing the stock, measured in number of items added during the period considered (the 9 weeks of the Summer Term, 1968).
OILL	Obtaining inter-library loans, number of items.
LML	Providing library materials for consultation in the library, number of user-hours spent using library material.
LL	Issuing items on long loan (2-week recall), number of items issued.
SL	Issuing items on short loan (4-hour or overnight), number of items issued.

US *Senior librarians giving advice to users*, number of hours spent on this activity by senior librarians.

UJ *Junior librarians helping users*, number of hours spent by junior librarians on this activity.

For other problems a different or a more detailed classification of activities may be appropriate.

The resources which must be taken into consideration include all those which limit the expansion of library activities. One very peculiar omission from the list of resources is the bookstock. This is because we are considering as activities the various ways in which use can be made of the existing bookstock, combined with one activity for keeping the bookstock up to date. The list of resources is given below.

S *Senior librarians*, man-minutes available for the listed activities during the time period under consideration (9 weeks).

J *Junior librarians*, man-minutes.

C *Clerical staff*, man-minutes.

P *Porters*, man-minutes

The budget available for purchasing books, in shillings.

Seats, seat-hours, being the number of seats times the hours that the library is open.

Shelving, spare shelving, in feet.

The resources required per unit of each activity and the total resources available are given in Table 3, which is similar in lay-out to Table 1.

Table 3

	I	OILL	LML	LL	SL	US	UJ	Resource available
S	72	4	0	0.6	0	60	0	2208 x 60
J	18	72	0.4	5	3	0	60	2364 x 60
C	18	0	0	0	0	0	0	460 x 60
P	2	12	0.1	0.1	0	0	0	233 x 60
Budget	56	5	0	0	0	0	0	4315 x 20
Seats	0	0	1	0	1	0	0	83600
Shelving	0.1	0	0	0	0	0	0	14000

FILMED FROM BEST AVAILABLE COPY

Table 3 yields most of the information needed to calculate the marginal costs. In practice only the budget and the various types of labour can be varied in the medium run. A decision markedly to increase the number of seats or shelves involves building a new library and is therefore a long-term decision. Since the budget is already measured in money terms, only the different hourly labour costs need to be found in order to calculate the marginal costs of the different activities. However, Table 3 does not give even an approximate picture of the production possibilities of the library. In mathematical terms, each column represents a variable; the rows could represent an equation, but only if all the resource is used up during the period. In practice the last two rows are not equations but inequalities. This shows us that the problem is not yet properly defined.

There are a number of additional constraints which in practice are limiting the production possibilities of the library. The various resources may be more limiting than our figures show. Senior librarians may be more specialised than our classification has allowed: some are cataloguers, some specialise in user-services, and it may be difficult to change from job to job. If all the shelving were to be used up during this period there would have to be a new library at the beginning of the next period if new stock were to continue to be brought in; but the amount of shelving available during the present period depends on the investment plans of the university, which depend in turn on the metaphysics of the U.G.C.

Apart from the resource constraints there are also demand constraints. The library may be able to provide more of certain activities but the relatively small population of users may not want such activities beyond a certain level. In an institution which engages in teaching and research the fact that users do not want certain activities beyond a certain level is not the same as saying that the marginal benefit of these activities is zero, because the marginal benefit is an assessment made on behalf of the institution as a whole and not by individual users for themselves. Most of us feel that everyone else should use the library more.

In order to postpone consideration of these difficulties but also to convey the impression that certain activities of the library are so important that they must be made available in all circumstances, we impose three demand constraints on the problem. These state that the number of long loans (LL) and short loans (SL), and the number of hours spent by users using library material in the library (LML) must be greater than or equal to the levels which actually occurred during the period. This completes the set of production possibilities.

Note that our problem at the moment is to find measures of the relative importance of library activities. To do this the actual levels of these activities are data. It is only after we have these measures that we can plan ahead to find the effects of changing the levels. Therefore it is not tautologous to take what actually occurred and deduce from this the measures of relative importance.

Table 4 gives in the first row the marginal costs in shillings of the activities. The second row gives the levels of the activities that would be predicted by using the marginal costs as the weights in the profit function (these weights are used in the same way that the prices were used in Part I); this is a calculation performed by the computer. The third row gives the observed levels that the library did produce.

Table 4

	I	OILL	LML	LL	SL	US	UT
Marginal cost	89.4	17.8	.064	.896	.40	24	8.8
Predicted level	1533	209	57,567	15000	1000	203	57
Actual level	1533	244	53,700	15000	1000	205	40

As a first approximation this is reasonable. It seems that an intuitive idea of marginal costs does influence librarians in the way they allocate resources. Unfortunately we were particularly interested in comparing the values of additions to stock (I) with inter-library loans (OILL) and the difference between predicted and actual is marked here.

Table 5 shows the weights on the profit function which result in the calculations yielding the same levels of activities as the observed levels. In the second row of Table 5 these same weights are expressed in terms of the value which is placed on one hour of a senior librarian's time. The ratios of the first and second rows are the same, and either row would yield the same result when used in the calculations.

Usually people find that a numerical approach to values is made easier by thinking in terms of swap-rates. The third row of Table 5 expresses the information of the other rows in terms of the number of units of each activity that the organisation is prepared to swap for one addition to the stock. For example, the figures under I and US imply that if the university were presented with a choice between one addition to stock and four hours of skilled user-services, it would choose the reader services; if the choice were between one addition to stock and three hours of user-services it would choose the addition to stock; a choice between one addition and 3.7 hours would leave the university indifferent. The comparison between I and LML will be particularly interesting to the teacher.

Table 5

	I	OILL	LML	LL	SL	US	UJ
Weighting	89.6	17.8	.064	.896	.40	24	8.8
Value, taking one unit of I as unity	3.7	0.74	.0027	.037	.017	1	.37
Swap-rate in terms of one unit of I	1	5	1370	100	217	3.7	10

Table 5 gives different ways of comparing the values of the library's activities, values to the university which have been obtained by studying the actual results of the decisions taken by the complex system of committees. A first approximation to the figures was found by using the marginal costs of the activities, but the figures given in the table are not costs: they are figures showing the relative benefits of the activities. Strictly, we should state that the library is behaving as if these are the relative benefits of the activities. Certainly we should take any opportunity to check the conclusions with the decision-makers concerned. However, the close approximation obtained by using the marginal costs does imply that the library knows what it is doing. There is a strong hint here that the library is fairly efficient in the economic sense.

If the conclusion is simply that the library is economically efficient, of what use is this technique for estimating the effects of library decisions? In the next section the technique is used to show the implications of adopting an innovation. When there is a complicated interaction between many resources and many activities a quick method of calculating "best" or "approximately best" ways of allocating resources is helpful.

One important qualification to this conclusion about efficiency must be borne in mind. The library appears to be allocating those resources which can be varied in the medium run (one or two years) in such a way that particular resources are being used to maximum advantage. But this assumes that users are knowledgeable, and that is a dangerous assumption to make about an institution which engages in teaching and research. Indeed, such an assumption threatens the institution's raison d'etre. Perhaps a few man-hours devoted to informing users more fully about the library's stock would alter the benefit ratios considerably.

Changing the technology

At the moment of writing, Newcastle University is developing a method of automating the ordering and accessioning procedure. (inevitably, cataloguing and classification remain traditional.) The new method will reduce the amount of labour required for the activity "increasing the bookstock" (I). We guess that the amount of senior librarians' time needed will be reduced from 72 minutes per item to 60; junior librarians' time will be reduced from 18 minutes to 9; clerical workers' time will be reduced from 18 minutes to 6. What would be the effects of introducing this innovation?

The first reaction to such a proposed improvement in library technique is to assume that a greater increase in the bookstock will be possible than is the case at the moment. However, the library is not only constrained by labour but also by the budget available for buying the books. The production possibilities are similar to those shown in Figure 5 at the point e: removing one constraint on expansion leaves others which are equally constraining. The problem is to know which resources to expand in order to make the best use of the innovation.

One alternative to expanding the resources would be to increase the flexibility of the labour force by giving similar training to juniors and clerks. An easier method is to assume that hiring and firing of juniors, clerks and porters are relatively simple and to give to the librarian discretion to spend the budget on both books and non-graduate labour. This means that the budget is raised, but not the cost, to the university. Each activity that previously used up the resources of juniors', clerks' and porters' labour time now uses up a proportion of the budget, the precise proportion determined by the hourly cost of these different types of labour.

Table 6 sets out the proposed changes. In column one the present resource requirements for one unit of I are shown. In column two the requirements are shown if the innovation were to be adopted. Column three shows the new requirements when the different types of non-graduate labour are treated as a straight cost.

Table 6

	Present requirement	Post innovation requirement	Non-graduate labour as a money cost
Seniors' time	72	60	60
Juniors' time	18	9	
Clerks' time	18	6	
Porters' time	1.8	1.8	
Budget	55.6	55.6	57.4

Assuming that all the activities except increasing the bookstock (I), user-services by senior librarians (US) and by junior librarians (UJ) must be at least at the level they were at before, and assuming that the relative importance of the different activities stays the same (see Table 5), the new techniques yield the increases in levels of output as shown in Table 7.

Table 7

	I	OILL	LML	LL	SL	US	UJ
Before	1533	244	53,700	15000	1000	205	40
After	1610	248	53,700	15000	1000	430	0

The benefit from the innovation consists of an increase of 5.7% in additions to bookstock and more than double the amount of senior librarians' time available for user-services. Table 7 also illustrates that the present planning technique can be used only with considerable common sense, because the calculation results in the recommendation that there should be no user-services from junior librarians.

If the calculations show that a change of plan results in an obvious improvement, the blend of activities recommended provides a convenient method of planning the allocation of staff within the library. A new version of Table 3 is prepared which takes into account the changed technology, and staff are allocated according to the blend of activities which are to be produced.

Summary.

1. Find the set of production possibilities of the library. Since the library is the major supplier of information services, some demand constraints will be necessary in addition to the resource constraints.

2. Use the cost of those resources which can be varied over a one or two year period to calculate the marginal costs. Use the marginal costs of the activities as a first approximation to the weights in the "profit" function. Adjust these weights until the "profit" function, together with the set of production possibilities, yields a linear programming solution that is the same as the actual blend of activities produced during the period. (This adjustment is simple if the print-out contains information on upper and lower bounds to the weights.)

3. Translate these weights into swap-rates and check with the decision-takers.

4. Use these weights to predict and compare the outcomes of possible changes in library policy and techniques.

Obviously this method of library planning has a number of disadvantages. It involves concepts with which many librarians will be unfamiliar. The numerical approach to comparative benefits, or the relative importance of library activities, implies an accuracy which is not there, and the value of approximation is often least appreciated by those who are unfamiliar with quantitative techniques.

However, a first attempt to find measures of relative importance is necessarily approximate and further experience can increase the accuracy of the estimates. The computer print-out (using Mathematical Programming System) contains a large amount of information additional to that mentioned in this paper, and this shows where the range of values is wide.

All the data required for this approach are needed anyway for elementary library management problems: the time taken by a particular type of labour to do a particular job, the cost of a unit of a particular activity, the amount of the various resources which are available.

The method allows library problems to be tackled one at a time, but without losing sight of the whole complex of library operations and the objectives of the library.

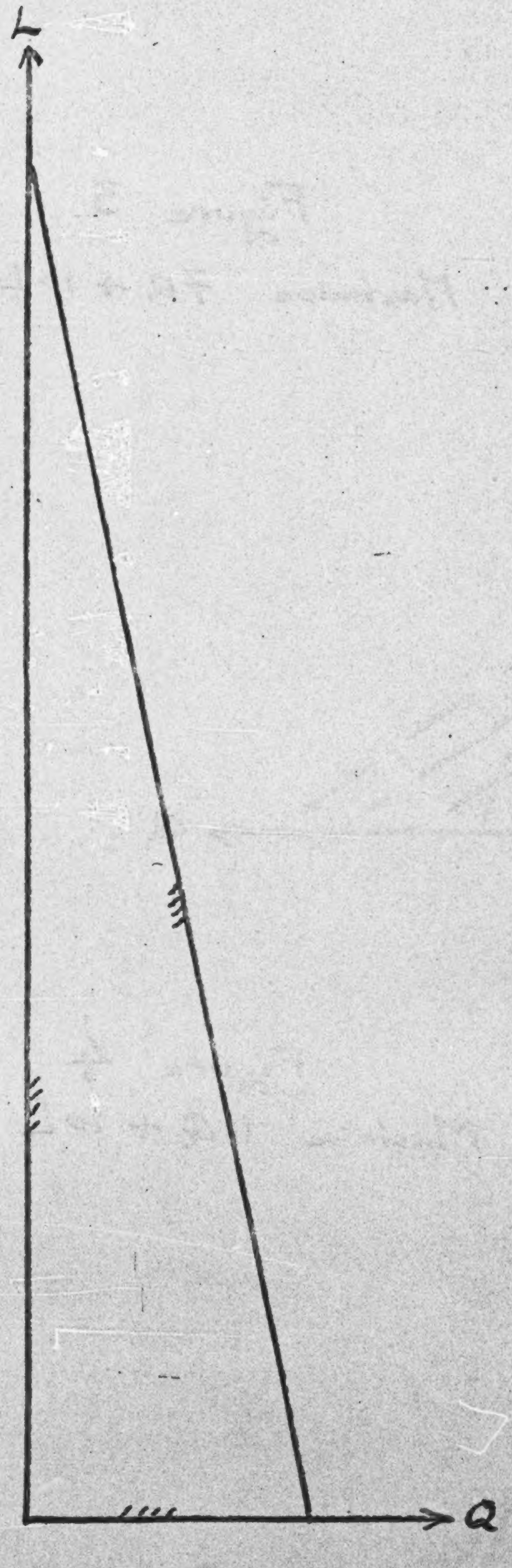


Figure 1

$11Q + 2L \leqq 980$

$Q \geqq 0$

$L \geqq 0$

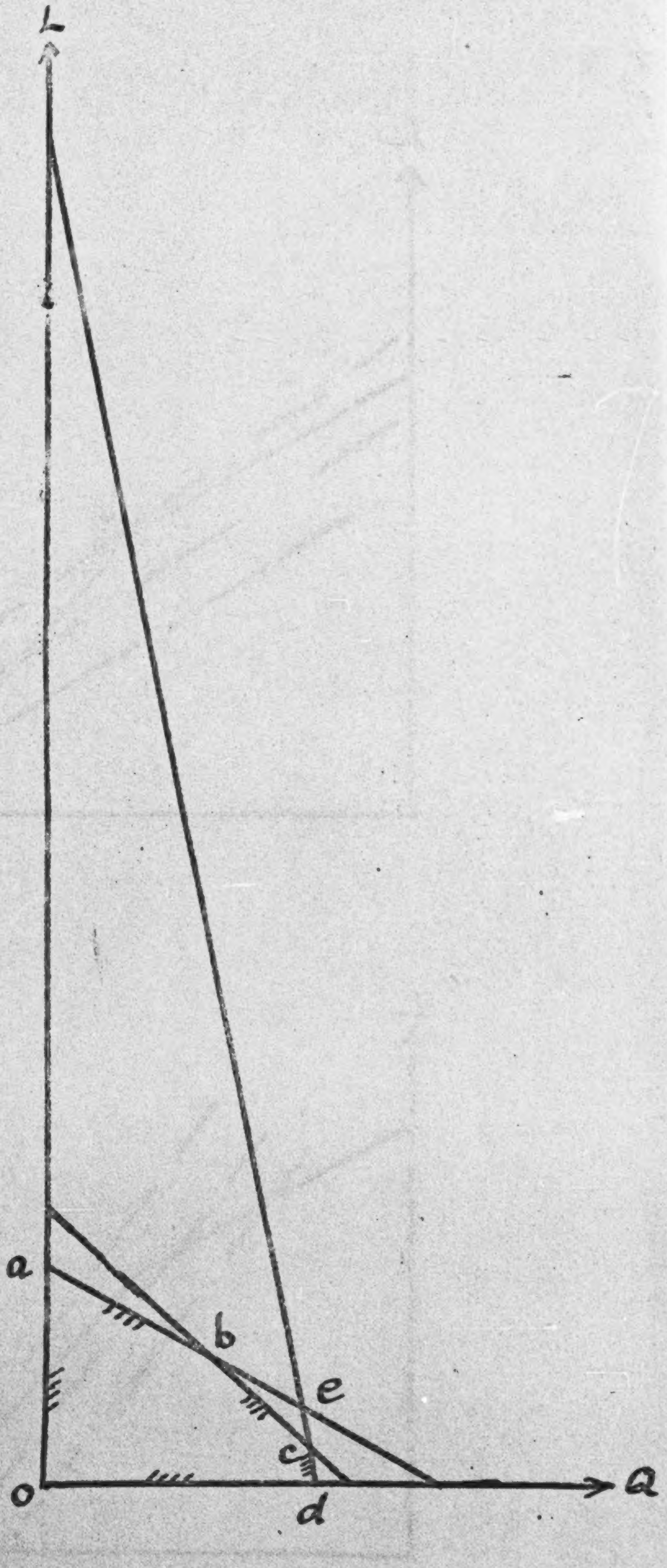


Figure 2

$11Q + 2L \leqq 980$

$3Q + 5L \leqq 400$

$Q + L \leqq 100$

$Q \geqq 0$

$L \geqq 0$

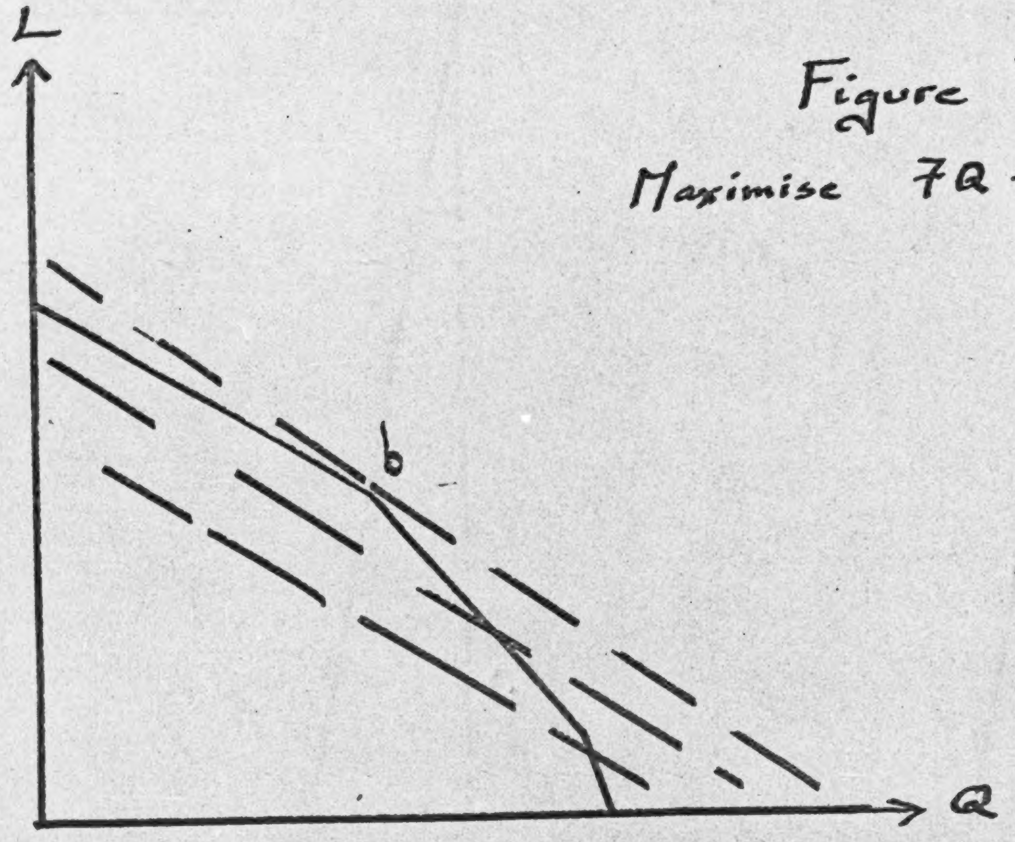


Figure 3
Maximise 7Q + 10L

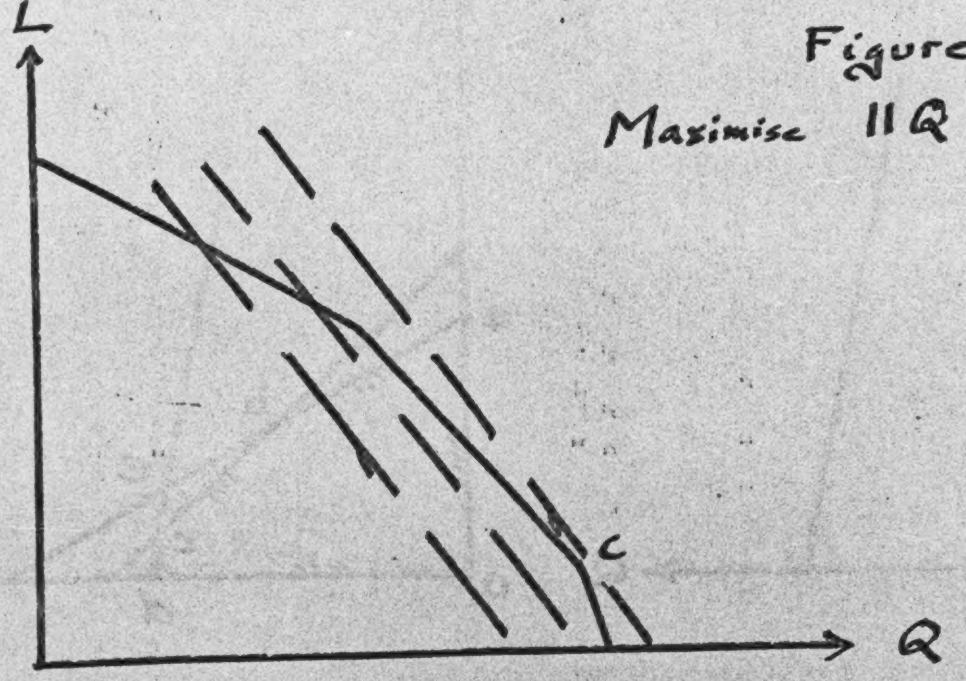


Figure 4
Maximise 11Q + 10L

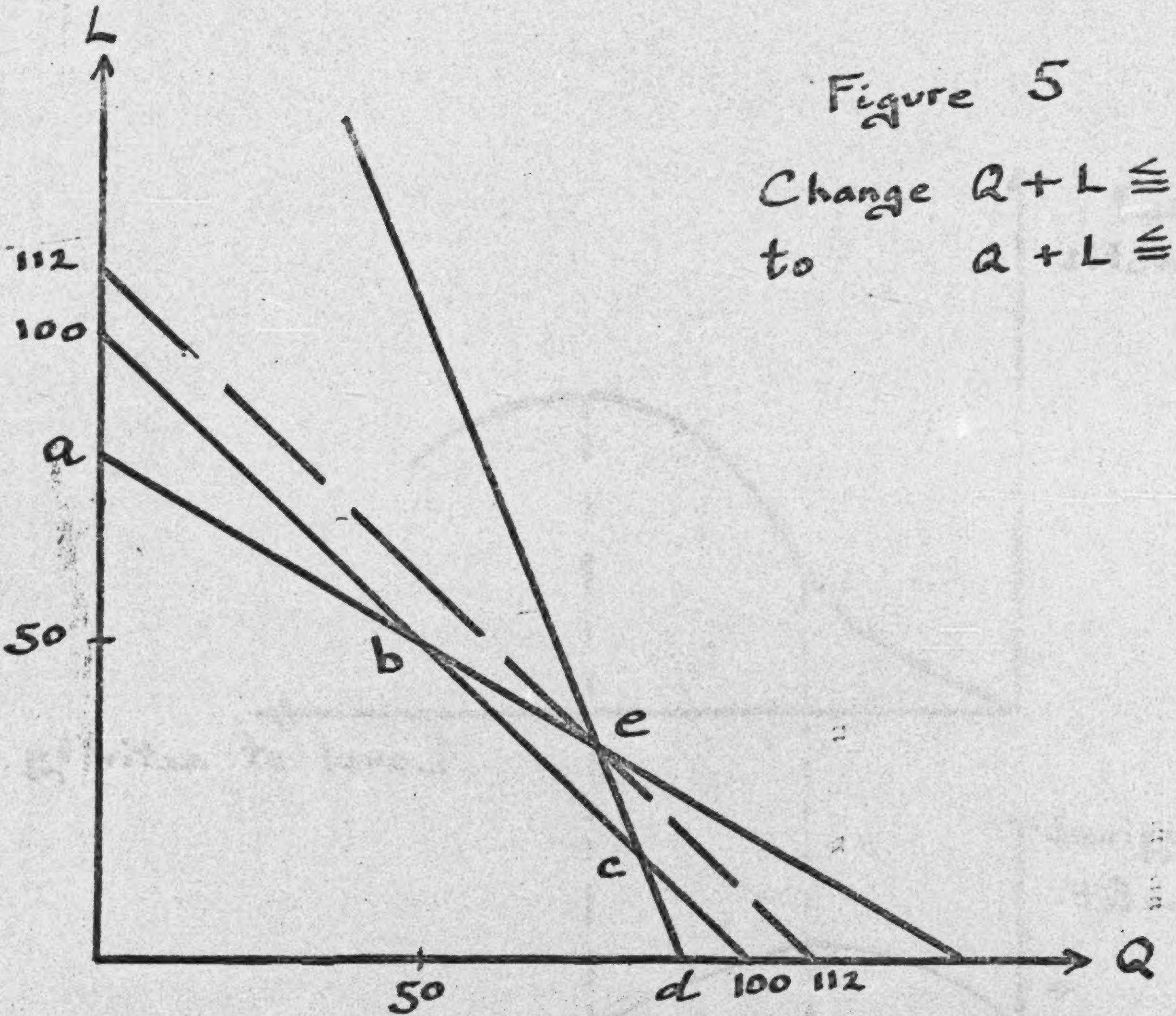


Figure 5

Change $Q + L \leqq 100$
to $Q + L \leqq 112$

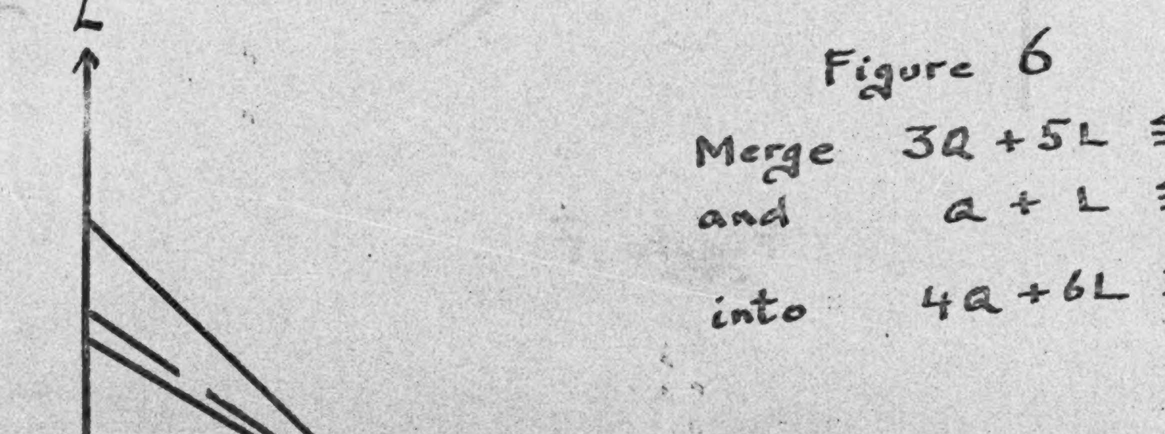


Figure 6

Merge $3Q + 5L \leqq 400$
and $Q + L \leqq 100$
into $4Q + 6L \leqq 500$

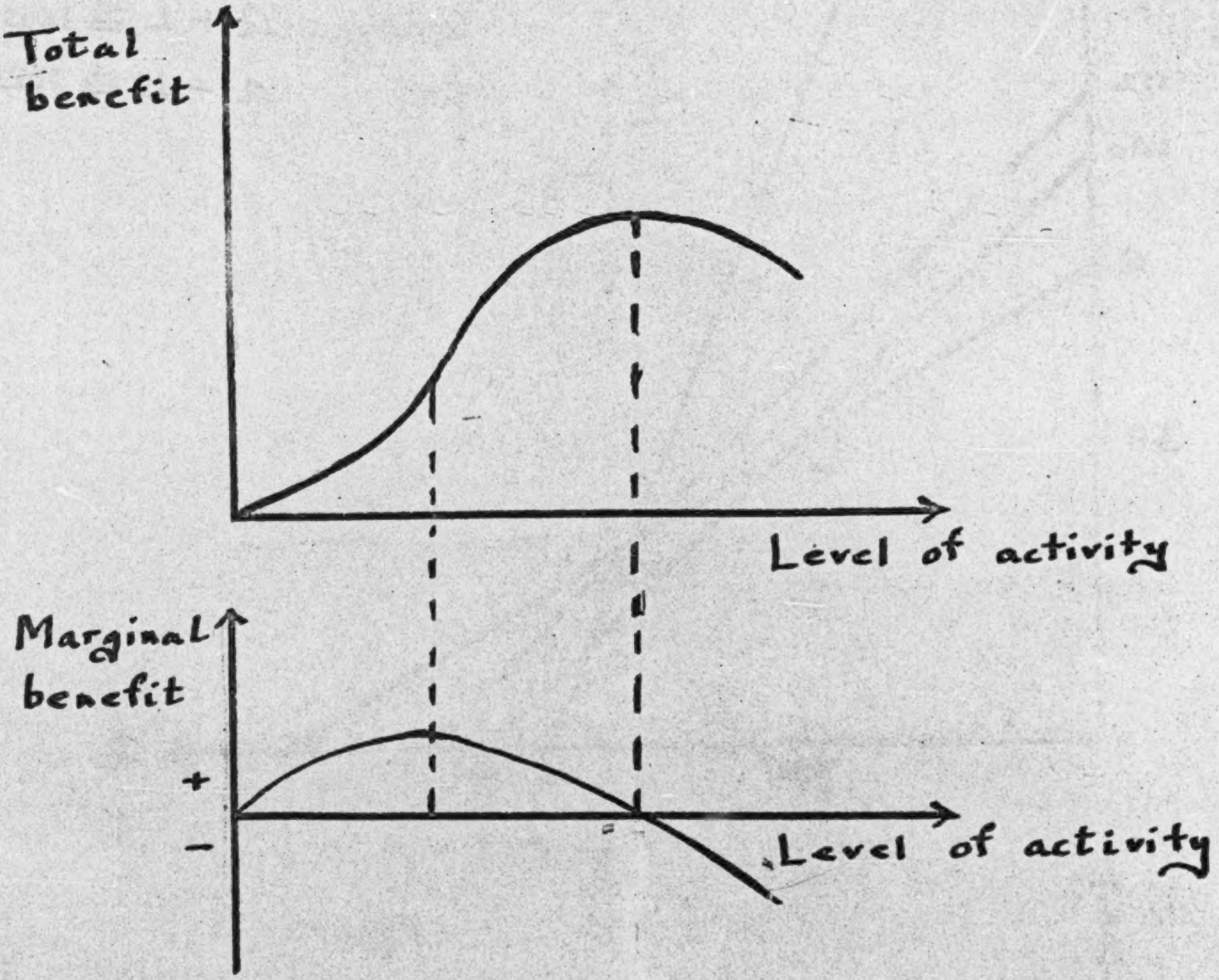

Total benefit
Level of activity
Marginal benefit
+
−
Level of activity

Figure 7

Discussion

Mr. Mackenzie said the thing which worried him about the Durham technique was that it was based on what the library offers rather than on the services it should provide; these were not necessarily coextensive. Mr. Morley replied that by using his technique and assuming that the library budget was fixed it was possible to vary certain objectives within the model and determine how the benefits would be affected.

Mr. Duchesne asked Mr. Morley why he had assumed that the short-run marginal and the short-run average costs were equal to the unit costs, and why they were assumed to be constant. Mr. Morley said this was only an assumption because it was not really known how these costs varied. Mr. Mackenzie observed that recent work had indicated that, for a variety of reasons, unit costs in the newer universities tended to be lower than in the established universities.

Dr. Urquhart suggested that one way to assess the benefit of a university library would be to find out what costs would be incurred by the users if a university did not have a library. More research was needed to find out what people did in the library; it might be possible for a university library to function with a farily low success rate, in terms of items being on the shelves when they are wanted, because if a user could immediately get some small proportion of his total demands he would be prepared to wait for the remainder to come as inter-library loans. However, there could be cases in literature searching when one book led to another, and it might be very costly if the first and key book were not on the shelves. Mr. Morley replied that it was important to think of the university library not only in terms of information retrieval, in which case the library responded to a stimulus, but as itself providing the stimulus by being available for browsing.

Mr. Carter said that Mr. Morley had been trying to establish how a university library could make the best use of fixed resources, but university administrators would like to know what proportion of university resources should be allocated to the library to give maximum benefit to the university. Mr. Morley replied that one way of doing this would be to assess the cost of alternative services which the readers would have to use, but which the university would have to pay for, if the services provided by the library were removed. One obvious result would be that travel allowances would have to be increased to pay for journeys to other libraries. If a range of such alternatives were examined and costed then some idea of the benefit provided by the library could be established.

Mr. Morley said that the Durham team had done a survey of the time students spent being formally taught and the time they spent in private study, either in the library or elsewhere. Preliminary results indicated that there was a certain critical level somewhere between twelve and fifteen hours a week; if the time spent on lectures was increased beyond this point a significant fall was noted in the time spent on private study of both kinds.

Dr. Urqhuart said that investigations should be carried out into where students got their books from; there must be a considerable variation from one university to another in the role played by the public library in providing curricular reading. In Cambridge the proportion was probably small, in Manchester and Birmingham quite large. Did the universities know this and approve of it? Mr. Longworth said that public libraries were used to the influx of students during university vacations, but he did not object to this, as students are legitimate public library users.

Referring to Mr. Morley's paper, Mr. Vickery asked for an illustration of how the use of the "maximising" approach yielded some insight into the preferences of library users. Mr. Morley replied that only a vague insight was obtained but they found, for example, that only 25% of the readers used in the catalogue.

Mr. Duchesne asked if it was legitimate to extrapolate from the snapshots of the library at various points in time to a dynamic situation, and hence to predict how the library would alter when certain variables were altered. Mr. Morley replied that it was possible to predict the general shape of the curves using the snapshot technique, and the measurements on which the technique depended could be repeated as often as required to obtain information about how the system was behaving.

Mr. Dammers pointed out that, when assessing the change in costs which would result with the introduction of computers, Mr. Morley had not included the cost of developing the system. He replied that the Newcastle programmes would be free and computer time was free, but he agreed that for a more realistic figure of the cost of the innovations these costs should have been included. Mr. May said that if Mr. Morley had applied this technique in industry to a particular section in a firm, then he might have been accused of sub-optimizing; he wondered if it would have been better to look at the university as a whole. Mr. Morley agreed that there was a danger of sub-optimizing, but he had really used a measuring technique and he had only used an optimization technique in his paper when considering changes in library technology.

Mr. Buckland asked Mr. Morley if he had considered the interactions between the various services provided by the library, and wondered if the effect on other services could be calculated if more resources were put into one particular service. Mr. Morley replied he had assumed that there were no complicated relationships between the services, but this did not prevent the services being altered and calculating the resulting changes in the benefit. Mr. Carter regretted that it was difficult to decide on the allocation of resources between books and equipment; some departments placed more emphasis on equipment than on books, and it was a pity that there was no guidance on the relative benefits of these.

SESSION 3: TECHNIQUES (Continued)

B. C. Brookes	Statistical distributions in documentation and library planning
F. F. Leimkuhler	Stograge policies for information systems
I. Woodburn	A mathematical model of a hierarchical library system
Discussion	

Statistical distributions in documentation and library planning. B.C. Brookes.

Introduction

The application of statistical techniques to scientific documentation is motivated by the following objectives:-

1. To identify the statistical distributions that arise in various aspects of documentation because such identification is highly informative per se.

2. To seek empirical laws, correlations and other relationships between differing processes within documentation work when these processes have been statistically described.

3. To apply any empirical relationships so discovered to make more rational or economic use of documentation by special libraries and in the planning of information services.

4. To develop a theory of information and documentation processes. The search for a theory is not simply an end in itself but is the only means of giving intellectual coherence to a subject which at present is little more than a rag-bag of practical techniques and empirically established procedures. Any theory with the hope of unifying information "science" and its applications to library planning will have to be essentially statistical. But effective library planning depends on a better statistical picture of books, periodicals and their usage than we have at present.

At the present time, the needs for statistical data and the usefulness of statistical techniques are inadequately appredciated by documentalists and so most of the data available are incomplete or in some respects untidy; almost every set of data that reaches me has to be re-sorted by tedious clerical work. With the increasing use of the computer information services, statistical data should become more freely available once their value to development and rationalization has been demonstrated.

For the most part, the statistical distributions that arise in library and documentation work are already well-known - the binomial, the Normal or gaussian, the Poisson, the gamma and their variants - and the techniques used in analysing these standard distributions - such as the χ^2, the t and other significance tests - are already being applied to some extent. The important aspect of all this work is to understand the conditions under which particular distributions are likely to arise and to recognize the situations which call for particular tests. But such understanding and recognition requires some apprenticeship in statistical work - there are no short cuts to the acquiring of the skills needed.

In addition, there are some distributions which are peculiar to documentation work and are therefore of practical interest to documentalists. When they are well enough established, these distributions will also attract the interest of theoretical statisticians and we can then hope to enlist outside help in developing the coherent theory and subject needs.

This paper is therefore restricted to a very brief survey of occurrences of the standard distributions that have been noticed in library and documentation work together with a more detailed discussion of one of the distributions peculiar to documentation - the Bradford-Zipf distribution and some of its applications, especially to the planning of special library services.

Occurrences of some standard distributions in documentation.

1. The Normal (or Gaussian) and related distributions

The variate of the Normal probability distribution has the mathematical form

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \quad \exp -(x-\mu)^2/2\sigma^2 \quad , \quad (-\infty < x < \infty) \; . \; . \; . \; . \; . \; . \; . \; (1)$$

and its "shape" is shown in Fig. 1.

Though the variate x ranges theoretically from $-\infty$ to $+\infty$, almost the total area under the graph, symmetrical about its mean, lies within ± 3 of the mean value μ. The parameter σ is a measure of the dispersion or spread of x about its value. The smaller the value of σ (the standard deviation), the more concentrated and compac the distribution becomes.

The Normal distribution occurs very widely - so widely that those who discovered its importance to biology in the late 19th century a little too readily assumed that i was a natural normal law of Nature. So by naming it the Normal distribution they wron implied that quantities which did not conform with it were in some sense abnormal. Mo anthropometric data (the heights of men, the sizes of their feet, etc.) are normally distributed i.e. most of them are 'near the average' value and men of exceptional heig or exceptionally large (or small) feet are relatively rare, so ready-made clothing and foot-ware manufacturers know very exactly the relative numbers of different sizes they can hope to sell.

The Normal distribution arises when the variate x can be regarded as having a basic constant value, μ, which is subject to independent random errors each of which is equally likely to increase or to decrease μ by some relatively small amount. It is of great theoretical as well as practical importance because it is both the ultimat ancestor and ultimate descendent of many other distributions. It shows great stabilit of form: sums or differences of normal variates are also normally distributed and it has many other properties of practical use and theoretical interest.

Though the mathematical function of (1) is in some respects rather intractable (i has no simple integral), simple techniques of handling it are well-established and tab of the "normal probability function" or of the closely related "error function" are readily accessible. The simplest way of checking whether a likely-looking one-humped symmetrical distribution is in fact Normal is to plot its cumulative sum on "arithmeti probability paper". The distribution is Normal if, and only if, the resulting plot i a straight line, or nearly so. (Fig. 2.)

An occurrence of the Normal distribution that surprised me arose when following up a clue dropped by John Swets[1] and his proposed measure of I.R. effectiveness. Whe I plotted Cleverdon's[2] Aslib/Cranfield results[1] for recall and fall-out on arithmeti probability paper they fell fairly closely along straight lines when the variate was Cleverdon's level of coordination. (Fig. 3.) For me the implication is that I.R. is a process subject to Normal (or Gaussian) noise and that it should be possible to dev a theory of I.R., and of indexing, based on this fact. In this case the basic "quant ity" subject to Normal error is the indexing language description of the contents of document. There is therefore an inherent "noise level" in every indexing language wh can probably be measured by the degree of overlap in the Swets model of the distributions of recall and fall-out. However, I do not propose to pursue this analysis furt until I have access to a computerized I.R. system for experimental work.

Distributions directly related to the Normal arise if it is some function of the variate x rather than x itself which varies Normally. One of these, the log-normal distribution may become of some importance in documentation. In this distribution it is log x (which can never be a negative quantity) which is Normally distributed. Its equation is

$$p(x) = \frac{1}{kx\sqrt{2\pi}} \quad \exp -(\log x - m)^2/2k^2 \quad , \quad (0 \leq x < \infty).$$

and its shape is shown in Fig. 4.

This distribution arises also in economics, biology, nuclear physics and the sizing of dust-particles and so techniques for handling it directly are already well-established. If the cumulative sum is plotted on probability paper which has a log scale for the variate x, the plot is linear if x is 'log-normal'. When Vickery[2] published his data (J.Doc.) on the distribution of numbers of papers in scientific periodicals based on a 1/10 sample of the periodicals held by the N.L.L., I plotted the cumulative sum on log probability paper. The plot shows that the distribution is approximately log-normal and so we now know the "shape" of the distribution of scientific papers among the periodicals. The distribution offers one way of describing statistically the totality of scientific periodical literature and it should therefore be possible to show that the "sum" of all possible scientific periodical bibliographies converges to this log-normal form.

In the log-normal distribution the variate x can also be regarded as having a basic constant value (its mean) which is subject to independent random errors, but in this case each of the disturbing factors multiplies the constant value by an "error" just less than or just greater than unity.

2. The Poisson and related distributions

The variate of the Normal and its related distributions is continuous i.e. it is a quantity which can take any value within the range of the distribution and which is usually measured in some way. But if we find ourselves counting we are concerned with a discrete distribution in which the variate takes only integral values such as 0, 1, 2, 3, . . . (There are exceptions to this statement: if the number of discrete values is high enough, it may be convenient, and accurate enough, to regard the variate as continuous as in the log-normal example used above.)

One of the most familiar discrete distributions is the Poisson:-

$$p(x) = e^{-a}\frac{a^x}{x} , \qquad (x = 0, 1, 2, 3, . . .)$$

This distribution has only one parameter, a, and it can be shown that the mean, $\mu = a$ and the s.d., $\sigma = \sqrt{a}$.

If "events" are randomly scattered in time or space then their temporal or spatial distribution conforms with the Poisson law. The classical example is von Borthiewicz's data on deaths in the Prussian army from the kicks of a horse - a random enough event which conforms closely with the Poisson law.

Random events occur in special library and documentation processes too. For example, the arrival of customers at the service desk of a library conforms with the Poisson law. The law has now become the basis of a mathematically sophisticated theory of queues from which, for example, the optimum number of service points for queues with different statistical laws of arrival and departure can be calculated. But one of the snags of using the Poisson law in library processes is that though the situations in which it is likely to be applicable are easily recognized, it cannot be applied until the value of a has been determined i.e. prediction based on Poisson analysis is liable to be rather post hoc unless the value of a is already known - in which case any practical queueing problems that arise have already been experienced and may already have been resolved. However, it would be useful, for example, in planning expansion of the number of service points required if the number of users were expected to increase - but essentially on the basis of past experience. But it must be stressed that confident statistical analysis of queueing problems demands mathematical competence.

3. The negative exponential and related distributions

If events occur at random in time or along a line, the intervals between these events, a continuous variate, are distributed according to the negative exponential law (Fig. 5):-

$$p(t) = \frac{1}{a}\ e^{-t/a} , \qquad (0 \leq t < \infty) . . . . (1)$$

This law arises in the ageing of periodical literature and will be used in a later section of this paper. In the context of ageing the law is a simplification, possibly an over-simplification, of a law of a more general form of which the law above is the simplest member. The more general form, the gamma distribution has the mathematical form:-

$$p(t) = \frac{1}{\Gamma(n)} \quad t^{n-1} \quad e^{-t} \quad , \qquad (0 \leq t < \infty) \; . \; . \; . \; . \; . \; . \; . \; (2)$$

When t is small, the factor t^n dominates: as t increases, the negative exponential factor gradually takes over and the value of p(t) declines towards zero. (Fig. 6)

The above distributions are widely used in reliability theory in which the life of mass-produced components conforms with variants of (2). If (1) is applicable, reliability theory yields a result for the life of a multiple component which consists of n elemental components working in parallel: the multiple component continues to operate until the last of the elemental components fails. The mean time to failure of the multiple component is increased by a factor which depends on n:-

$$1 + \tfrac{1}{2} + \tfrac{1}{3} + \; . \; . \; . \; . \; . \; . \; \frac{1}{n}$$

This factor represents a law of diminishing returns: if n is 2, the factor is 1.50; if n is 3, the factor is 1.83; and so on. But the intriguing aspect of the factor is that it is also a Zipf distribution and so suggests that periodical literature which conforms with this law (discussed in a later section) also offers the user a law of diminishing returns as the size of any bibliography he is interested in increases.

The practical handling of these exponential distributions is usually helped by the use of semi-log graph paper in which the log scale is used for the ordinate p(t).

Statistical distributions in general

The distributions mentioned above have been grouped in a way which might suggest that they form independent entities. But in fact they form a closely inter-related family with an underlying web of relationships. For example, the Poisson and Normal laws can both be derived from the Binomial law: and the Poisson law reverts to the Normal law as a increases indefinitely. One advantage of developing the statistics of documentation is that this web of underlying statistical relationships can help to link documentation and library usage phenomena which are still regarded as unrelated.

The Bradford-Zipf law of scatter of periodical literature

If all the issues of all the periodicals which publish papers relevant to a well-defined scientific topic are collected together for a period of, say, one year, the number of relevant papers published by each contributing periodical can be counted. When the count is known, the periodicals can be arranged in order of decreasing productivity. The periodical ranked 1 may well be one which is almost wholly devoted to the particular subject or it could be a more general scientific periodical, such as Nature or Science published frequently and containing large numbers of short articles and letters. In all subjects there will be further periodicals which contain substantial numbers of papers relevant to the subject. But there will also be a long tail of periodicals - about half the total number of contributing periodicals, which only contribute on average between one and ~~two relevant~~ papers per annum. The cumulative totals, $R(n)$, of the numbers of relevant papers are then found. Thus $R(1)$ is the total of relevant papers in the periodical ranked 1: $R(2)$ is the total of relevant papers in the periodicals ranked 1 and 2 and $R(n)$ is the total of relevant papers in the first n of the ranked periodicals.

When successive values of $R(n)$ are plotted against n on semi-logarithmic graph paper, the graph will be found to begin with a rising curve which runs, sooner or later, into a surprisingly smooth straight line. (Fig. 6.) The useful feature of this "bibliograph" is that if enough data have been collected to ensure that the straight line is reached, the end of the line can be predicted. And if the end of the line is known, so also is the total number of periodicals which contribute at least one paper per annum to the scientific subject. The narrower or more specialist the subject, the shorter will be the curve which rises to the starting point of the line; in many cases it suffices to identify and count the productivities of only the 10 most highly productive periodicals.

The equation of the straight line is $R(n) = N \log n/s$ where N is the total number of periodicals contributing at least one paper per annum to the subject[3]. To estimate the value of N it is necessary only to find where the straight line, if continued downwards, meets the axis of n. The highest (and therefore the most reliable) values of $R(n)$ and n are then substituted in the above equation. We then have

$$N = \frac{R(n)}{\log n/s}$$

In this equation the logarithm is the 'natural' or Napierian logarithm which has the base e: tables are found in most elementary sets of 4-figure mathematical tables.

When N is known, the total number of relevant papers published per annum can also be calculated since

$$R(N) = N \log N/s .$$

The above formulae are based on a mathematical formulation of Bradford's law of scatter[3]. They represent a mathematical ideal which is very closely approached by computer searches of comprehensive bibliographic collections, such as MEDLARS, (Fig. 7) but manual searches (Fig. 8) are likely to fall short of the Bradford ideal. Enough evidence has accumulated to suggest that deviations from the Bradford ideal can be accounted for by admitted or demonstrable incompleteness or by weaknesses in the search procedures. The Bradford ideal can therefore be taken as the practicable target when sufficient care and thoroughness over the search is taken.

In many practical cases, it is neither necessary nor would it be economic to collect the "complete" Bradford bibliography. However, it is helpful to know what fraction of the total is attainable or practicable. For example, if anyone were interested in establishing a special library for a research department interested only in muscle-fibres, the bibliograph shows that of the contributing periodicals listed in the MEDLARS bibliography, the nuclear group of 25 periodicals alone produce 57% of the relevant papers and that 80% of the field is covered by the 64 most productive periodicals whereas it requires 313 periodicals to cover the "complete" bibliography.

Other applications of the bibliograph

A bibliograph of the same form as that obtained for the periodical literature of a well-defined scientific subject can be obtained from other types of bibliographic data. The Bradford-Zipf law seems to arise in situations in which a finite set of choices is available but in which the probability of the choice of a particular item is reinforced by each previous choice of that item so that a few of the items eventuall attract most of the choices. The "popularity" of the most frequently chosen items, however, eventually causes "saturation" - not all the first choices of these items can be satisfied. For example, a periodical may attract more papers than it has space to publish: the rejected papers may then be submitted to less popular periodicals in the ranked sequence until they are accepted. "Saturation" is indicated by the curved part of the graph.

Fig. 9 shows the bibliograph of abstracting publications drawn from the data of Wood and Bower[4] of the N.L.L. on the use of abstracting publications in social sciences These data conform very closely with Bradford-Zipf expectations. They show a nucleus of 10 publications and an expected total of about 80 of which 77 were included in the N.L.L. list.

The data on citations is more difficult to interpret. East and Weyman's[5] data on citations of books in the literature of plasma physics produce a bibliograph of the expected form except that it continues beyond the expected limit. The graph indicates a nuclear group of only 5 books and, if the Bradford-Zipf law were to hold for citation a total of 222 cited books would be expected. East and Weyman, however, found that 529 books were cited but they add that the most cited items were "clearly directly related to plasma physics" and that a "large number" of less cited monographs dealt with "other areas of physics, mathematics, etc." Clearly, the items of choice in this particular case, the books cited in the literature of plasma physics, do not all belong to the same bibliographic classification. The eleven books ranked 1 to 11 certainly belong to the literature of plasma physics and so account for the conformity of the initial part of the graph with the typical Bradford-Zipf form.

Data on the citations to periodicals in the same subject are also provided by East and Weyman (op.cit.). In this case the bibliograph of the 70 periodicals most frequently cited lacks the initial curve: the graph consists initially of two straigh lines of different slopes. (g. 10) The first straight line is a typical Zipf graph indicating that "saturation" has not occurred here. It is not surprising; there is no limit on the number of times a periodical may be cited though, oddly enough, "saturation" occurred in the data on monographs. The Bradford-Zipf estimate in this case also exceeds the number actually found (14,173 estimated, 11,901 found) whereas with books the estimate was less than the number found. A possible partial explanation of the anomalies in this case may stem from the fact that the citations noted were in no way limited by the date of the issue of the periodical cited. It is reasonable to assume that conformity with Bradford-Zipf expectations would require that all the "items" open to "choice" - here the periodicals cited - should be exposed to choice for the same period of time. This condition is clearly not satisfied by this particular set of citation data (there is no reason why it should have been) one of which is mention as going back as far as 1870.

A third set of data published by East and Weyman (op.cit.) enumerates the citatio for each paper cited from Physics of Fluids (1961-4). The bibliograph is shown in Fig. 11. This again is of the typical Bradford-Zipf form and again the straight line extends (248 calculated, 372 found) beyond the Bradford-Zipf expectation. Examination of these data suggests however that the most frequently cited paper (which is cited only 14 times) is not cited as frequently as a Bradford-Zipf distribution of the given totality would require nor is the range of "popularities", from 14 citations to 1, as large as would be required. It may be possible to regard such data, and a similar typ of data which arises from consideration of residual demands, such as those handled by the N.C.L. and the N.L.L. by some form of truncated Bradford-Zipf distribution from which the unknown "head" is missing. But the empirical evidence so far available suggests that, in such cases, the "tail" is not chopped off cleanly from the "head" and further analysis is needed.

These partial failures of the Bradford-Zipf bibliograph help to elucidate more precisely the situations in which the bibliograph can be used with confidence and to clarify the conditions which have to be satisfied. Citations clearly do not fully satisfy these conditions. There is no saturation effect because it 'costs' nothing, in effect, to cite a paper; it may be that the seeming conformity of sets of citations with the Bradford-Zipf form for sets of periodicals is merely a reflection, partly distorted, of the underlying periodical distribution. "The 'core' journals tend to attract the most citable papers", East and Weyman noted.

At the same time the distributions of citations widen the linear logarithmic form of cumulative sums of documentary distributions and so suggest that a more general empirical law may be found. In the Bradford-Zipf bibliograph for periodicals the value of s (the point at which the linearity intercepts the axis of log n) is never less than 1. But in Fig. 9 the initial linearity has a value of s, 1/2.08, which is just less than 1/2, while, for the second linearity, s is again 1.

One area to which the distribution should apply, since the required conditions seem to be satisfied, is to the issues of books from a library. Here again, the library user has a very wide choice and yet his choice of the most popular books is restricted because such books are not always available when he wants them and because the process of issuing and restoring a book to circulation takes finite time even in a computerized issue system. A bibliograph of the issue records of a library should therefore establish which books, at any one time, constitute the popular nucleus and should enable an estimate to be made of that fraction of the total collection which is of active use. But no book issue data have yet been studied in this way.

A second area to which the distribution in one of its forms should apply is to the frequency of use of index terms in a document retrieval system. Here again a bibliograph should be helpful in identifying the terms which are over-worked or under-worked, in devising labour-saving search strategies, or in estimating the ratio of growth of the admissible indexing terminology.

In all the areas mentioned more data and analyses are needed. One use of the bibliograph arises from its potentiality as an instrument for the managerial control of a library or documentation system.

Theoretical problems

Clearly some theoretical problems remain to be solved. If the points of the initial rising curve of the bibliograph are plotted on log-log paper instead of on semi-log paper, a single straight line is produced in the simplest cases. This indicates that the form of the initial curve of the bibliograph is

$$R(n) = \alpha n^{\beta}$$

where α and β are constants with $\beta < 1$. But sometimes a series of two or three successive straight lines arises, in which the value of β increases in steps up towards 1. In all such cases so far studied, these complex forms arise when the total bibliography covers a long time-span or when two or more slightly different categories of data are known to have been combined. Fortunately, the effects of these variations affect only the nuclear contributions, which have to be counted anyhow, and they do not affect the practical applications based on the subsequent linearity.

The Bradford-Zipf distribution also has an extraordinarily stable form. It is therefore possible to calculate the composing or resolving of wholly independent or partially overlapping bibliographies. It would also be helpful to know exactly how the form of a bibliograph develops as the literature on a new subject grows from the first basic paper; usually, it seems, a subject inevitably widens as its literature grows and the resulting bibliograph may be regarded as the envelope of a number of Bradford-Zipf linearities of increasing slope.

Statistical Bibliography

Empirical laws such as the Bradford-Zipf could be helpful in transforming library management into a more business-like activity. The study of bibliography is still in a kind of natural history phase, much concerned with particulars but little concerned as yet with seeking general relationships between the particulars. The fact that the emergent empirical laws of bibliography cannot yet be adequately explained in causal terms need not prevent their application in solving appropriate problems. There is much work for them to do in the planning of more rational, economic and effective library and information systems.

In such work the Bradford-Zipf distribution has a role to play because it has an optimizing factor built into it; once periodicals are arranged in order of their productivities of relevant documents or books in order of popularity, any cut-off left at the 'head' of the bibliograph automatically indicated the n best periodicals, or the n most popular books. It should therefore be helpful in designing systems or when applied to them. There follow three simplified models intended to show some applications of the Bradford-Zipf distribution.

1. Complete periodical sets and the buying of photocopies

Problem: The complete set of periodicals which produce papers relevant to a scientific topic T numbers N. This set contains a long tail of periodicals ($\approx \frac{1}{2}N$) which produce an average of only 1 to 2 relevant papers per annum. Consider the economics of buying photocopies of relevant papers instead of periodicals.

Let A = average cost of the periodical subscriptions per annum (in any monetary units);
P = average cost of one photocopied paper (in the same monetary units);
and $P/A = p$, $(0 < p \leqslant 1)$.

It would be reasonable to assume that a special library interested in the subject T would, in any case, wish to buy all the periodicals in the Bradford nuclear zone of the set so that the possibility of acquiring photocopies applied only to periodicals in the outer zones. The analysis can therefore be based on the periodicals which conform with the Zipf linearity i.e. it will be assumed that $n > c$ throughout.

Let the B-Z formula of the complete set be

$$R(n) = N \log n/s \quad , \qquad (c \leq n \leq N) \;.\;.\;.\;.\;.\;.\;.\;. (1)$$

The productivity of the periodical ranked nth will then be N/n relevant papers per annum and the cost of each of these relevant papers will be An/N. Thus the cost of buying relevant papers by means of periodicals increases uniformly with n but the cost of buying photocopies remains constant at P units each. (Fig. 10.) The point at which it would pay to switch to photocopies is given by

$$An/N > P \;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;. (2)$$

i.e. when the cost of buying relevant papers through the periodicals exceeds the cost of a photocopy. The point is reached when

$$n > PN/A = pN \;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;. (3)$$

The total cost of buying the first n periodicals is nA. The number of photocopies to be bought is equal to the number of relevant papers to be expected per annum from the periodicals ranked (n + 1) to N inclusive. This number is

$$\begin{aligned} R(N) - R(n+1) &= N \log N/s - N \log (n+1)/s \\ &= N \log N/(n+1) \;.\;.\;.\;.\;.\;. (4) \end{aligned}$$

and their cost will be

$$PN \log N/(n+1) \;.\;.\;.\;.\;.\;.\;.\;.\;.\;. (5)$$

The total cost of buying the complete annual bibliography will therefore be

$$C = nA + PN \log N/(n+1) \;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;. (6)$$

a function of n which has a minimum value at n + 1 = PN/A. Substituting this value of (n + 1) into (6), we have

$$C = PN - A + PN \log A/P \;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;. (7)$$

As the total cost of buying all N periodicals would be NA units per annum, the saving, S, per annum is

$$NA - PN + A - PN \log A/P$$

so that

$$S/AN = 1 - \frac{P}{A}(1 + \log A/P) + 1/N$$

$$(1 - p) + p \log p \;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;.\;. (8)$$

The graph of this function, which represents the maximum proportional (or percentage) saving attainable is shown in Fig. 1[illegible]. The economy achieved is clearly closely dependent on the value of p = A/P.

Example: If N = 400, s = 2, the total bibliography would consist of 400 log 200 = 2119 papers per annum with a nucleus of 6 or more key periodicals.

Assuming that p = 0.2, the number of periodical subscriptions could be reduced from 400 to 80. The number of relevant papers thus acquired would be 400 log 40 = 1476 and the remaining 643 papers, distributed over 320 periodicals, would be bought as photocopies. The library would thus have complete periodical average of its subject at 52% of the cost of buying the complete set of periodicals.

2. The discarding of ageing periodicals from a subject set

Problem: The library takes only M of the N periodicals of the complete subject set. Policy: Periodical issues are discarded when their "utility" has decayed to the same fixed level. Consider the loss of utility.

Application of this policy would leave the library with relatively long runs of the most productive periodicals and relatively short runs of the least productive, with a minimum level of utility applied to all the periodicals.

The Mth periodical produces N/M papers per annum. If this periodical is discarded after time T, the lost utility is

$$\frac{N}{M} e^{-T/a} = k \quad , \text{ (say).}$$

And so

$$\frac{N}{M} e^{-T/a} = k = \frac{N}{n} e^{-t_n/a} \quad \ldots\ldots\ldots (1)$$

where t_n = age of discard of the periodical ranked n. Solving (1) for t_n we get

$$t_n = T + a \log M/n \quad , \quad (n \leqslant M) \ldots\ldots (2)$$

The total proportional loss of utility $= \dfrac{Mk}{N \log N/s}$ .... (3)

If the Mth periodical is to be retained at all, k cannot be greater than the value of N/M which it attains when T = 0. (In this case the Mth periodical is bought, examined for its occasional relevant papers which are noted, but it is not shelved; all other periodicals acquired are retained for periods which increase from zero upwards as the rank n decreases. Note that if T = 0, then k = N/M and (3) reduces to

$$\frac{1}{\log N/s} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

which is independent of M, the size of the collection taken. (Fig. 13.) Hence, if periodicals were removed completely one by one, beginning with the least productive, 1/N th of the total original utility would be lost at each step as the utility of the remaining periodicals is reduced to the same minimum level.

For N = 400, s = 2, log N/s = log 200 = 5.30. The adoption of this policy of discarding periodicals when, as they age, they reach the same level of utility, would mean discarding 1/5.30 or 19% of the total utility with a saving of about 60% of the shelf space.

3. The optimization of the periodical stock of a central/branch library system

Problem: A library system consists of a central library together with r branch libraries all serving the same subject for a total population of R users. It 'costs' the user of the system α monetary units for each consultation of the periodicals in his branch library and β monetary units ($\beta > \alpha$) in his central librar The central library is bibliographically "complete" in the subject, i.e. it takes all N relevant periodicals. If it is required to minimize the total cost (i.e. user cost plus provision cost) how many of the N periodicals should be made available at the branch libraries?

Let n $(1 \leqslant n \leqslant N)$ be the required number of periodicals. Then the proportion of consultations that can be satisfied at the branch libraries is

$$\frac{N \log n/s}{N \log N/s} \text{ and the remaining proportion, } \quad 1 - \frac{N \log n/s}{N \log N/s} = \frac{\log N/n}{\log N/s}$$

have to be referred to the central library. The cost of each of the R users is there

$$\alpha N \log n/s + \beta N \log N/n$$

The cost of providing n periodicals for each of the r branch libraries and N periodicals to the central library is

$$(rn + N)\, a$$

where a is the average cost of the periodicals. So the total cost of the system is

$$R(\alpha N \log n/s + \beta N \log N/n) + a(rn + N) \quad \ldots\ldots\ldots\ldots \text{(1)}$$

If n is regarded as the variable while all other terms remain constant, the minimum total cost occurs when

$$arn = NR(\beta - \alpha)$$

i.e. when $\dfrac{n}{N} = \dfrac{R}{r}\dfrac{(\beta - \alpha)}{a}$

It will be seen that n increases as the ratio of users to branch libraries increases or as the disparity between the costs of using the central and branch libraries increases. But the value of n is also inversely proportional to the average costs of the periodicals. These results accord with expectations.

The analysis could be taken farther. It has so far been assumed that the library system has its fixed number r of branch libraries but the above result would hold for any value of r. By substituting the value of n just formed into (1), the total cost is expressed as a formula from which the unknown n is now eliminated. The formula thus becomes an expression for r by means of which the optimum number of branch libraries could be calculated. It should be noted, however, that, if the number of branch libraries is theoretically allowed to vary, it cannot be assumed that α and β are constants independent of the value of r. It is likely that α decreases and that β increases as r increases but no helpful data on this problem are known to the author.

Concluding Comments

The unusual feature of the Bradford-Zipf distribution is that it is a concurrence of two distributions. How does this double distribution arise? This question of being analysed with the help of a computer simulation program which, at present, is only partly worked out, and by means of which we test our hypotheses. We have found a means of generating the Zipf distribution (the linear part of the bibliograph) from a simple probability specification which simulates random and unrestricted publishing of papers in a number of "potentially contributory" periodicals which grows as the subject develops. It has already become clear, however, that the probability mechanism which generates the Zipf distribution will have to be modified to generate the Bradford nucleus and that the modification requires restriction of the simulated publication of papers in the nuclear periodicals. Possible probabilistic forms of this restriction are now under test. Finally, the generating program will have automatically to permit a smooth relaxation from "restricted" to "free" publication at the critical point of the simulated bibliograph.

The new results reported in this paper are mainly theoretical. But, it is suggested, they are plausible enough to justify further exploration, both theoretical and practical. The Bradford-Zipf distribution and its bibliograph need further analysis which, ideally, requires access to a comprehensive computerized I.R. system to produce quick and reliable results of the kind and on the scale detailed analysis demands. The possible application of this new distribution to monographs, as well as to periodical literature, and also to library processes deserves to be explored because there is already some evidence that the Bradford-Zipf distribution has a much wider generality than was first appreciated. It could be very useful in planning any rationalization of library resources.

References

1. Brookes, B.C. The measure of information retrieval effectiveness proposed by Swets. J.Doc. 24,2. (1968). p. 41.

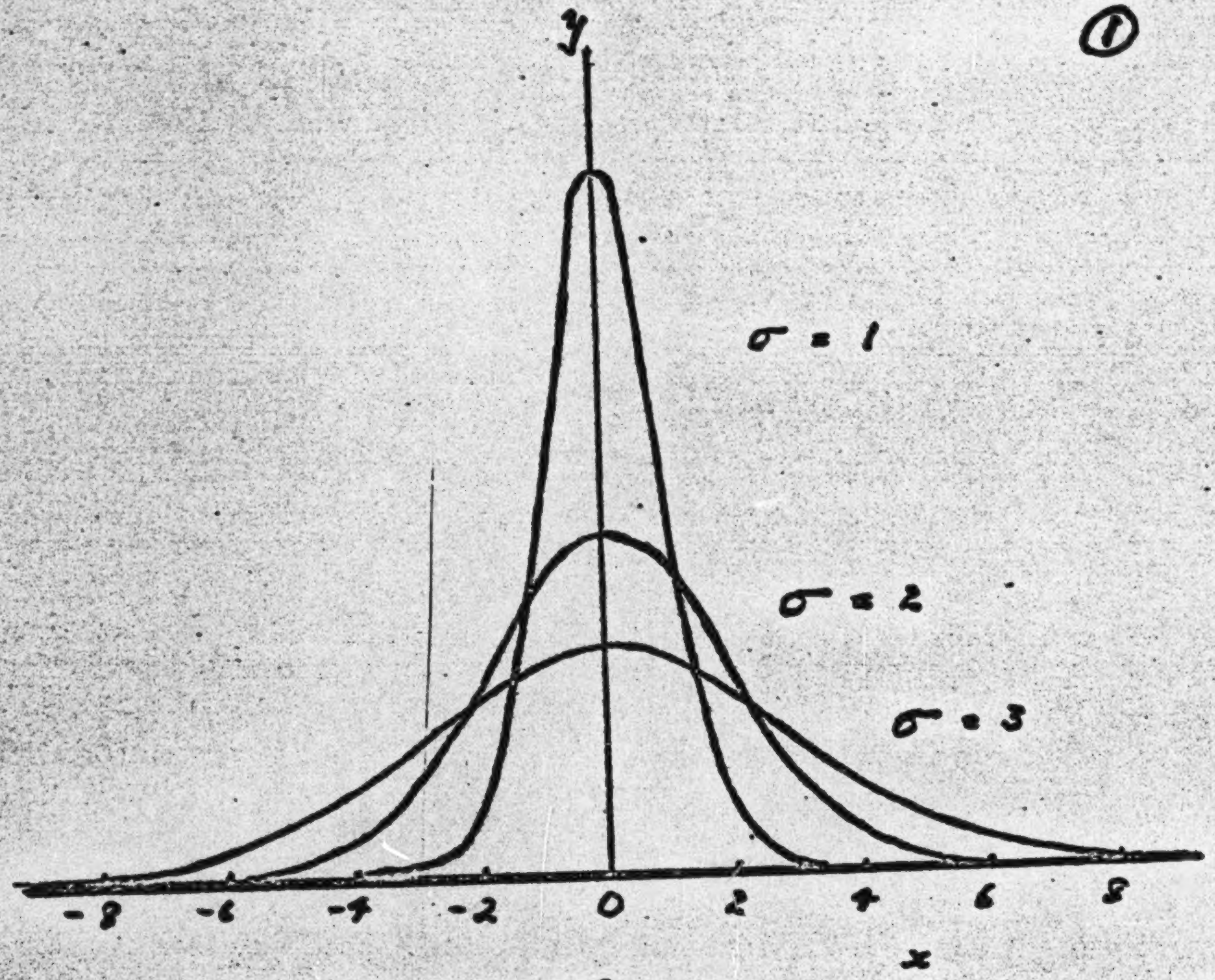
2. Vickery, B.C. Statistics of scientific and technical articles. J.Doc. 24,3. (1968). p. 192.

3. Brookes, B.C. The complete Bradford-Zipf 'bibliograph'. J.Doc. 25,1. (1969). p. 58.

4. Wood, D.N. and C.A. Bower. The use of social science periodical literature. J.Doc. 25,2. (1969). p. 108.

5. East, B. and A. Weyman. A study in the source literature of plasma physics. ASLIB Proc. 21,4. (1969). p. 160.

①



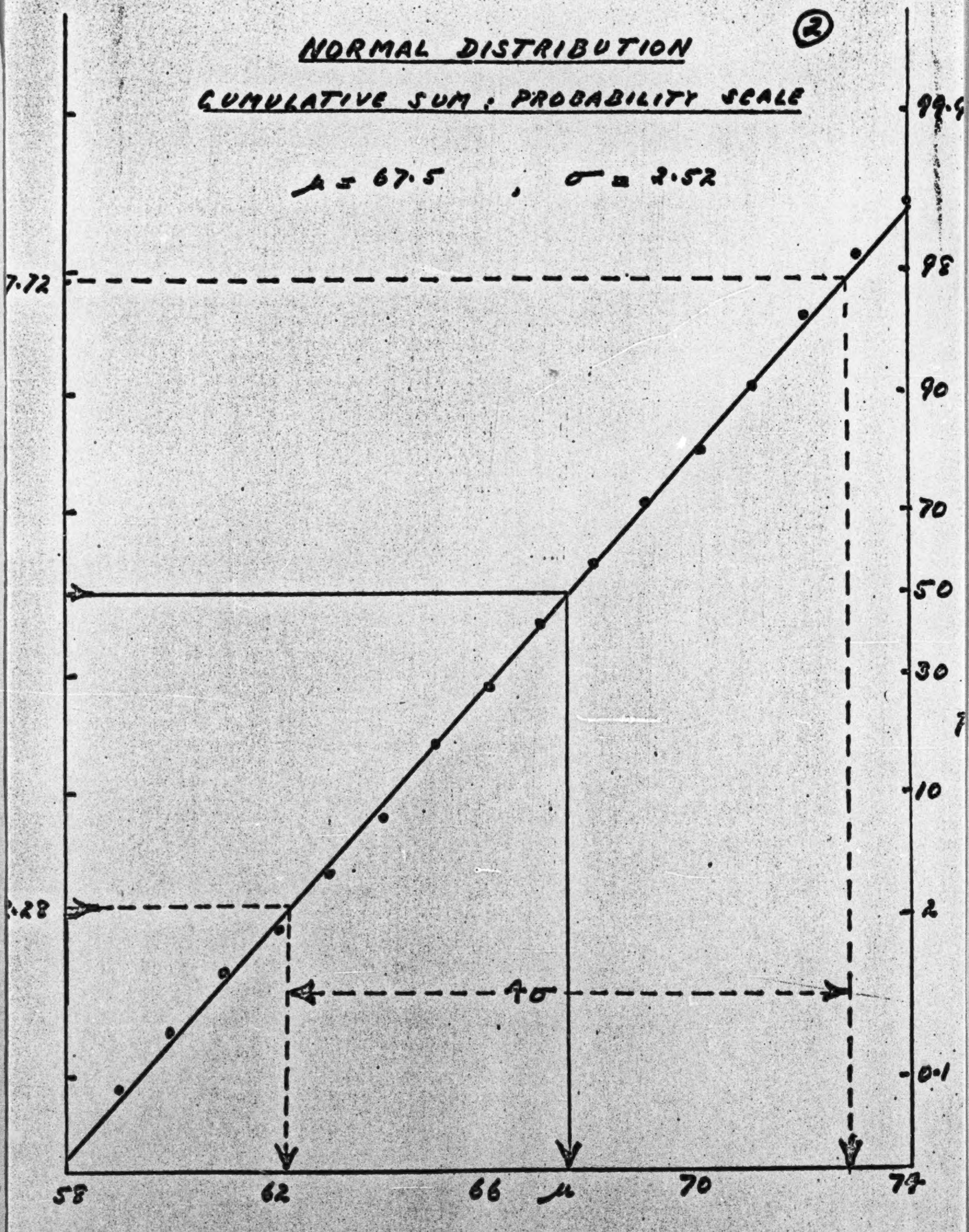
NORMAL CURVES

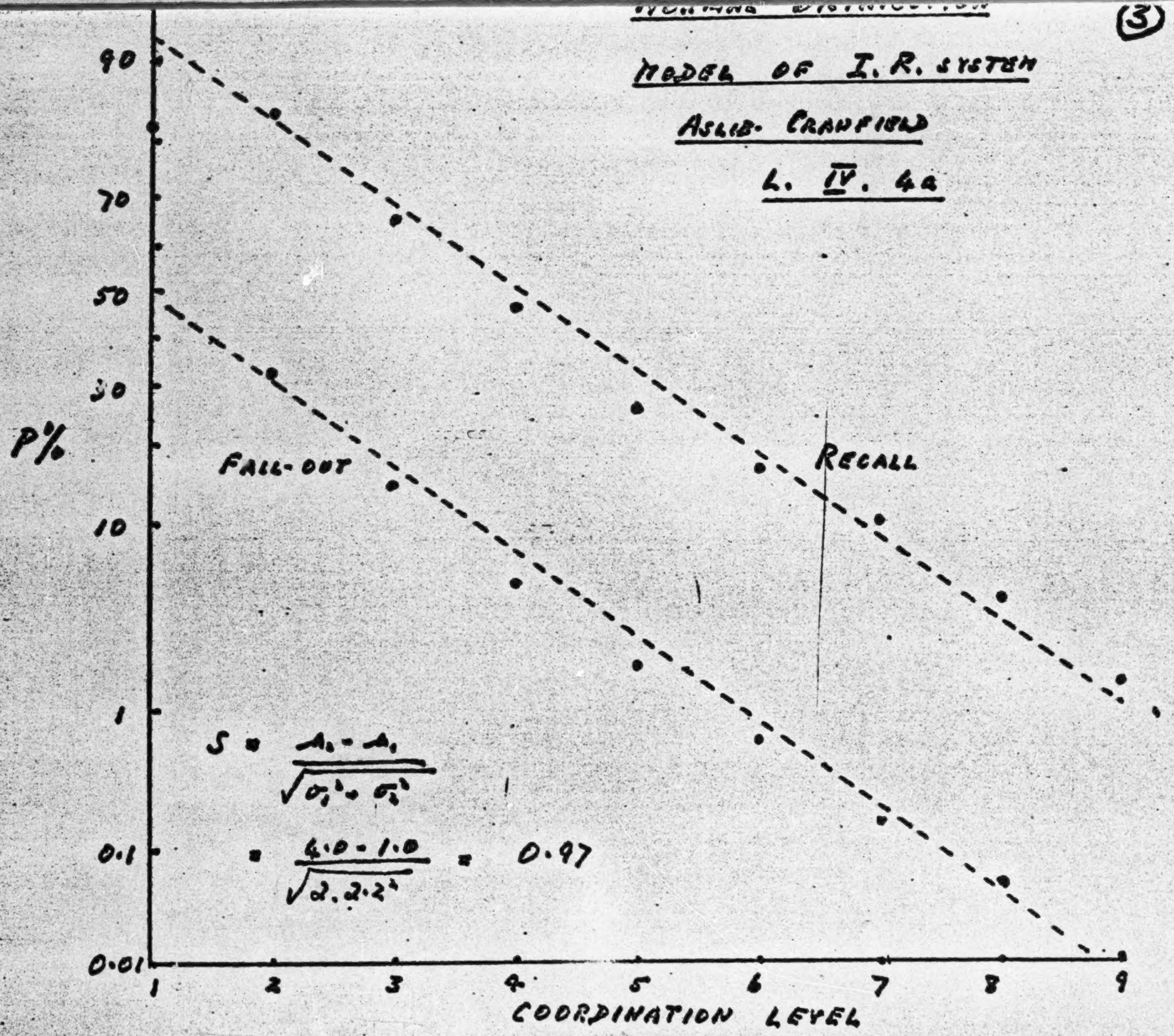
②

NORMAL DISTRIBUTION

CUMULATIVE SUM : PROBABILITY SCALE

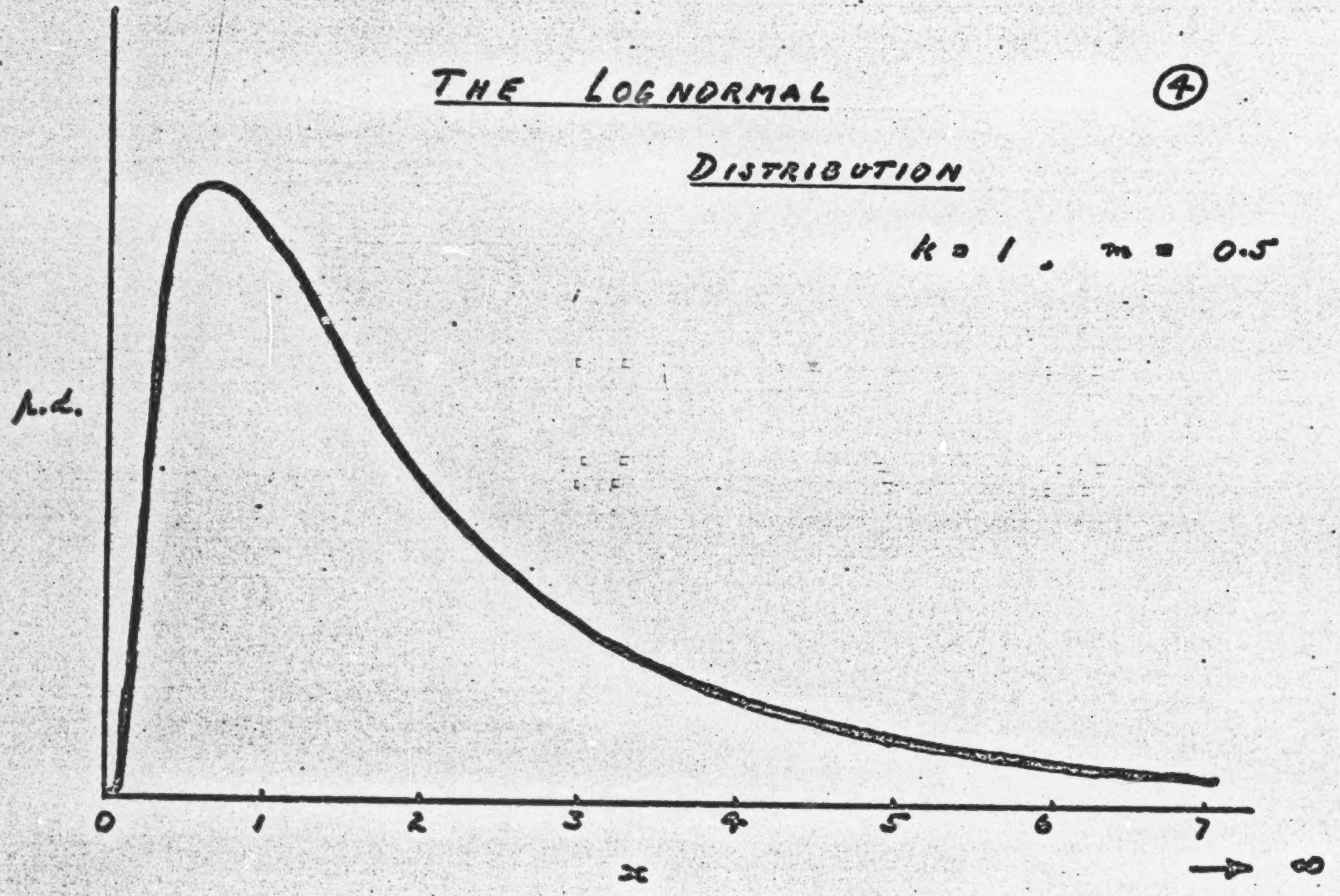
$\mu = 67.5$, $\sigma = 2.52$





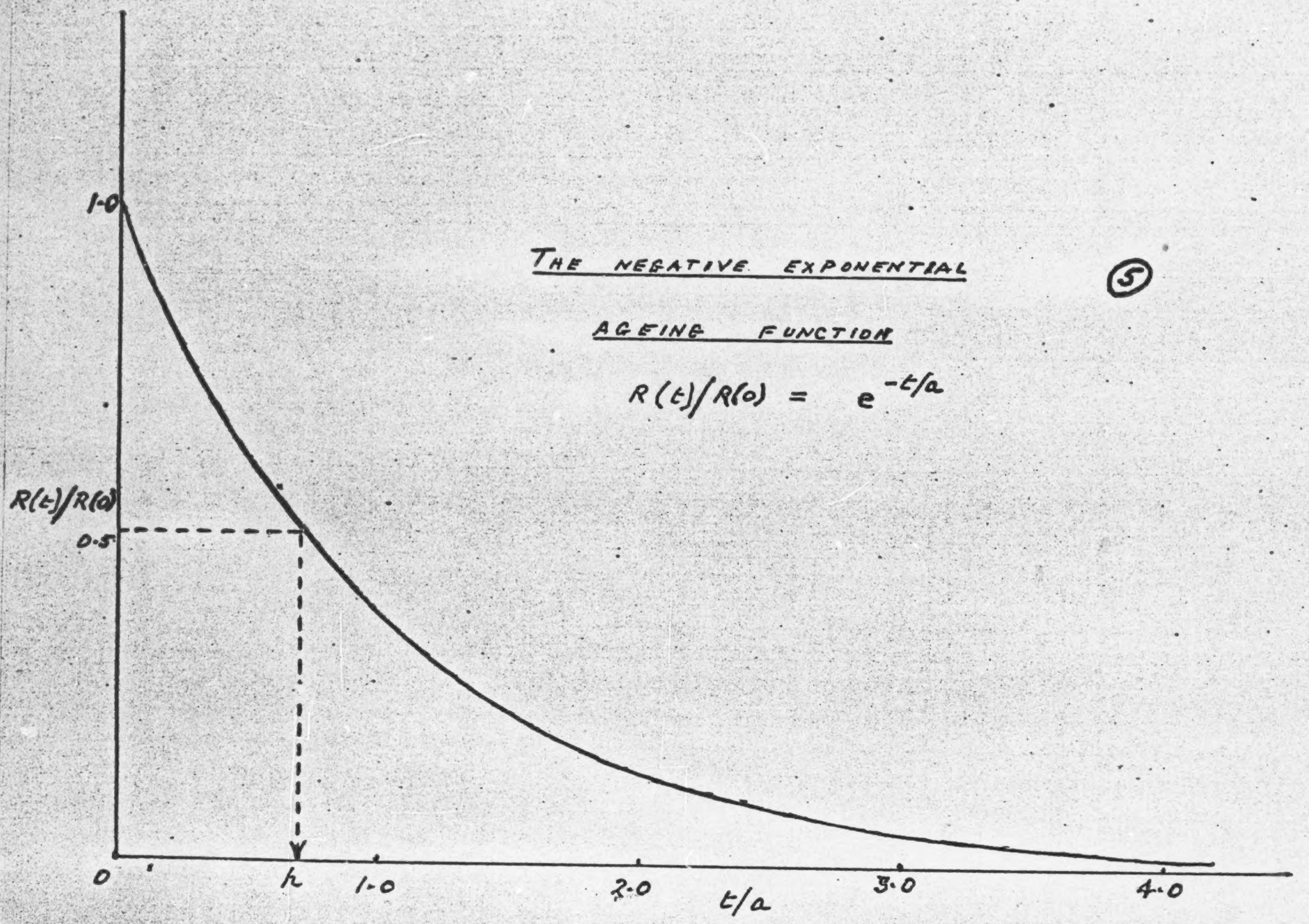
$$S = \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

$$= \frac{4.0 - 1.0}{\sqrt{2 \cdot 2.2^2}} = 0.97$$



THE LOGNORMAL DISTRIBUTION ④

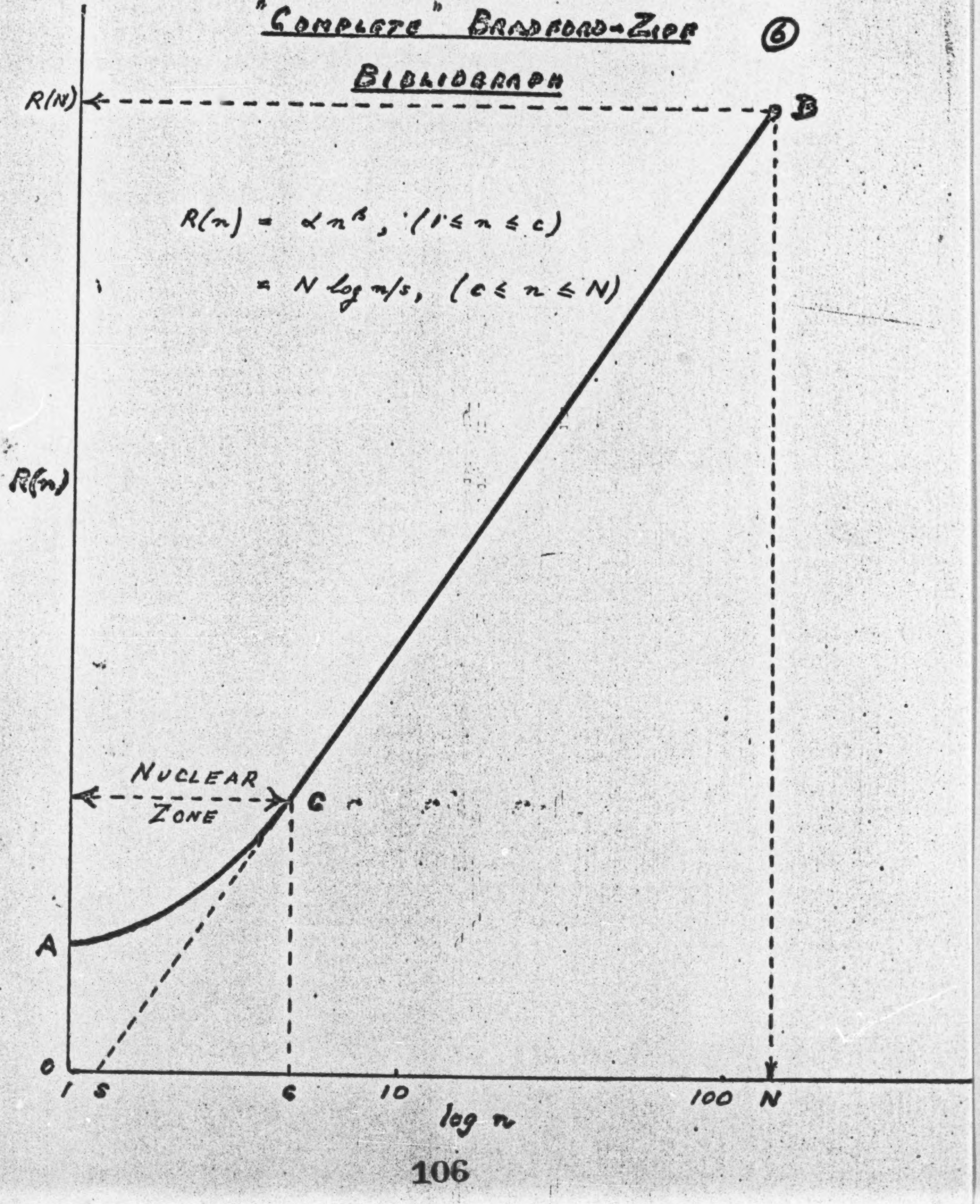
$k = 1, \ m = 0.5$



THE NEGATIVE EXPONENTIAL AGEING FUNCTION ⑤

$$R(t)/R(0) = e^{-t/a}$$

"COMPLETE" BRADFORD-ZIPF BIBLIOGRAPH ⑥



$$R(n) = \alpha n^{\beta}, \quad (1 \le n \le c)$$
$$= N \log n/s, \quad (c \le n \le N)$$

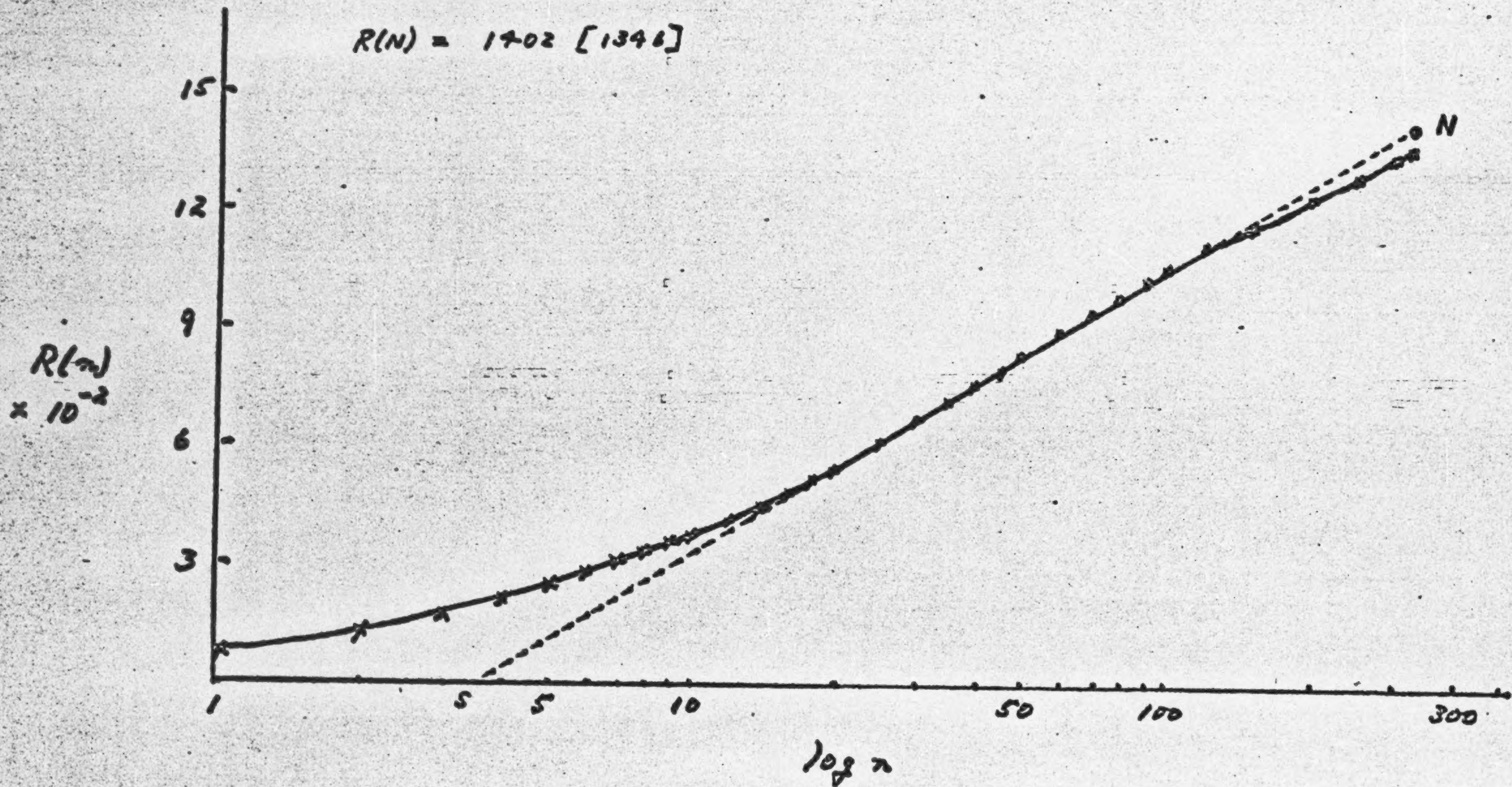
⑦

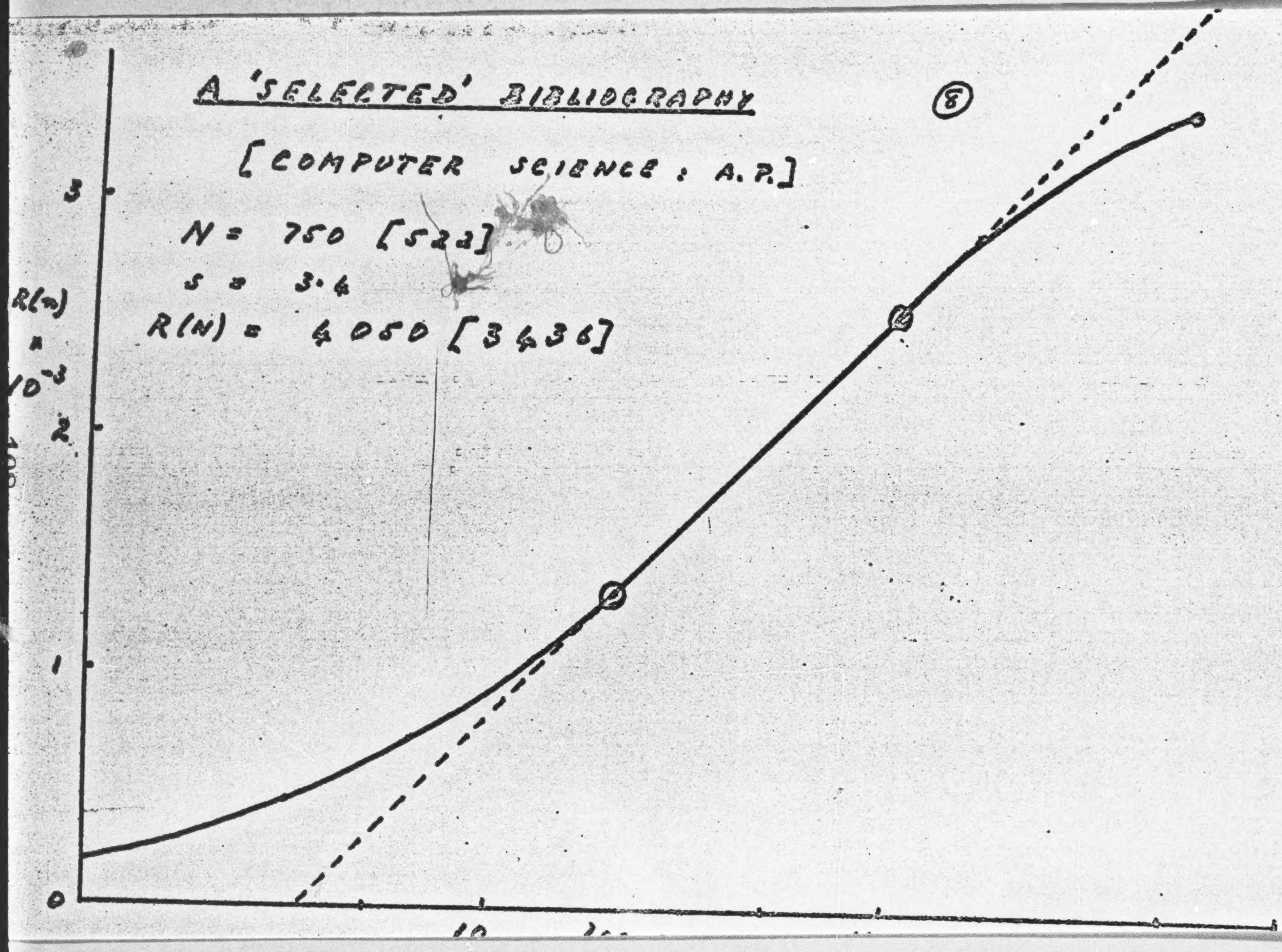
MUSCLE - FIBRE BIBLIOGRAPH

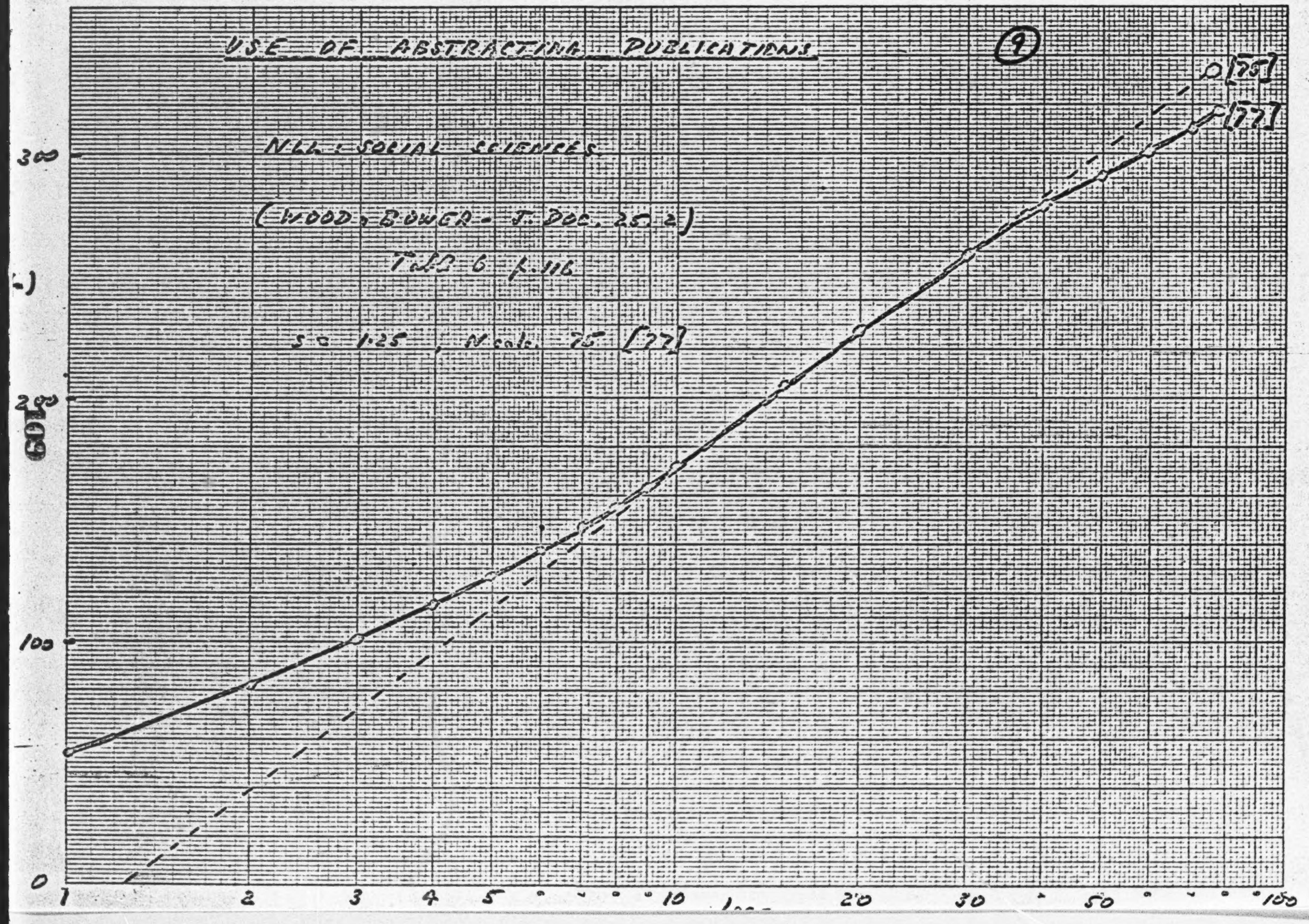
MEDLARS SEARCH

$N = 313$ [313]

$R(N) = 1402$ [1346]




A 'SELECTED' BIBLIOGRAPHY
⑧
[COMPUTER SCIENCE : A.P.]
N = 750 [522]
s = 3·4
R(N) = 4050 [3436]
R(n) × 10^-3
3
2
1
0


USE OF ABSTRACTING PUBLICATIONS
9
NLL: SOCIAL SCIENCES.
(WOOD, BOWER - J. DOC. 25 2)
TAB 6 p.116
S = 1·25 , N calc. 75 [77]
O [75]
[77]
300
200
100
0
1
2
3
4
5
10
20
30
50
100

CITATION DATA : PLASMA PHYSICS

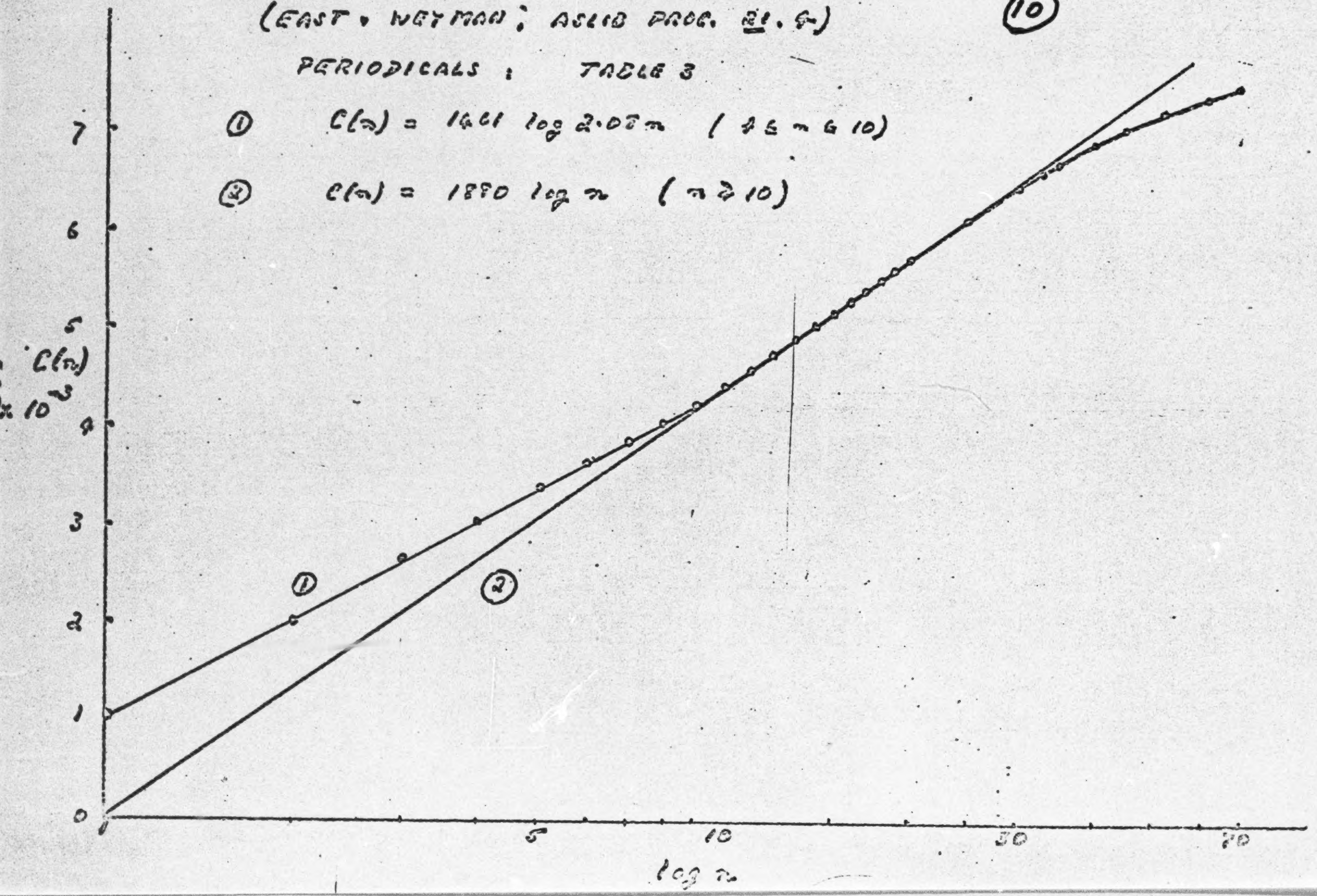
(EAST + WEYMAN ; ASLIB PROC. 21, 6)

PERIODICALS : TABLE 3

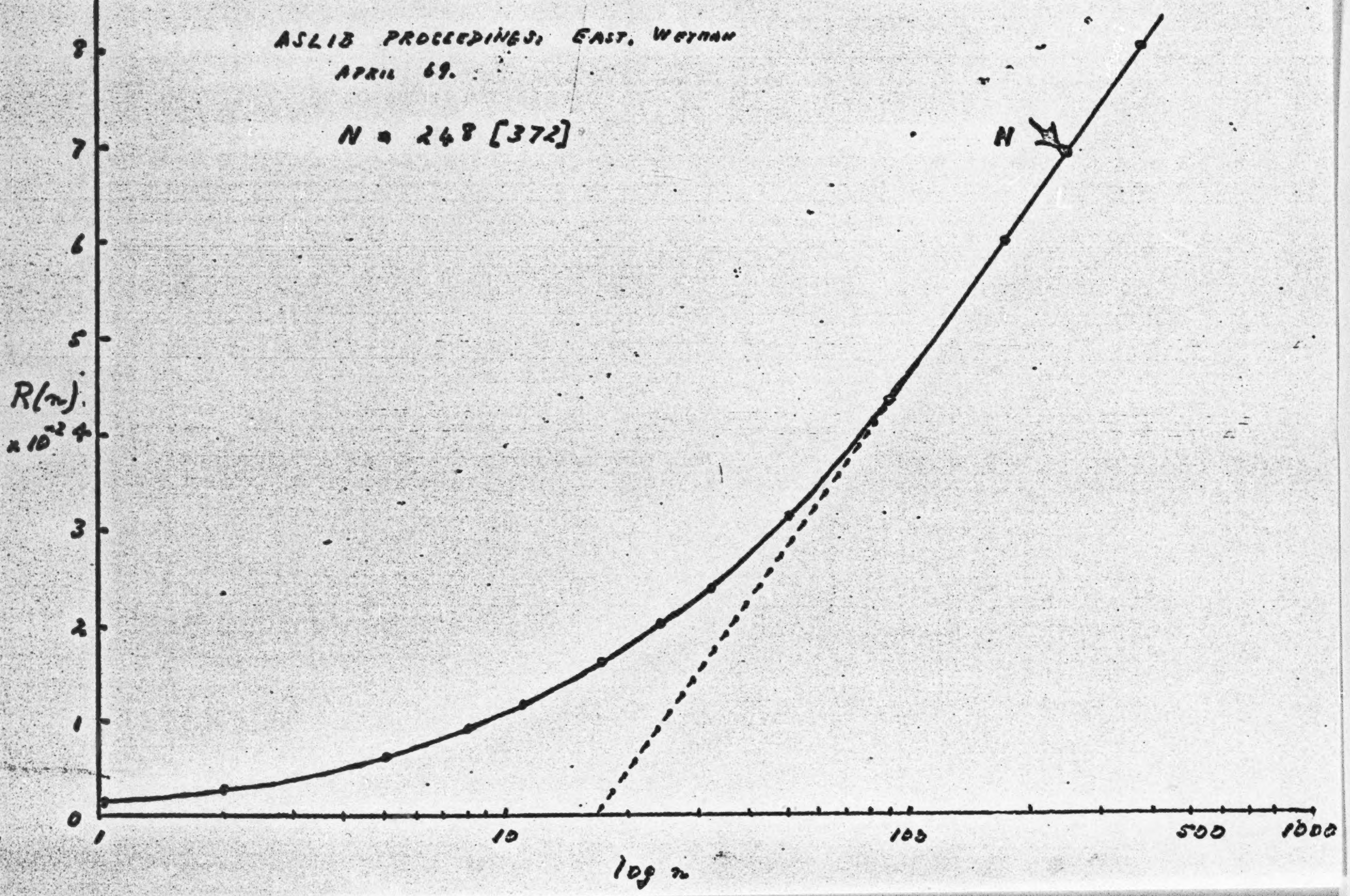
⑩

① $C(n) = 1461 \log 2.08n \quad (1 \leq n \leq 10)$

② $C(n) = 1880 \log n \quad (n \geq 10)$



CITATIONS : PHYSICS OF FLUIDS ⑪



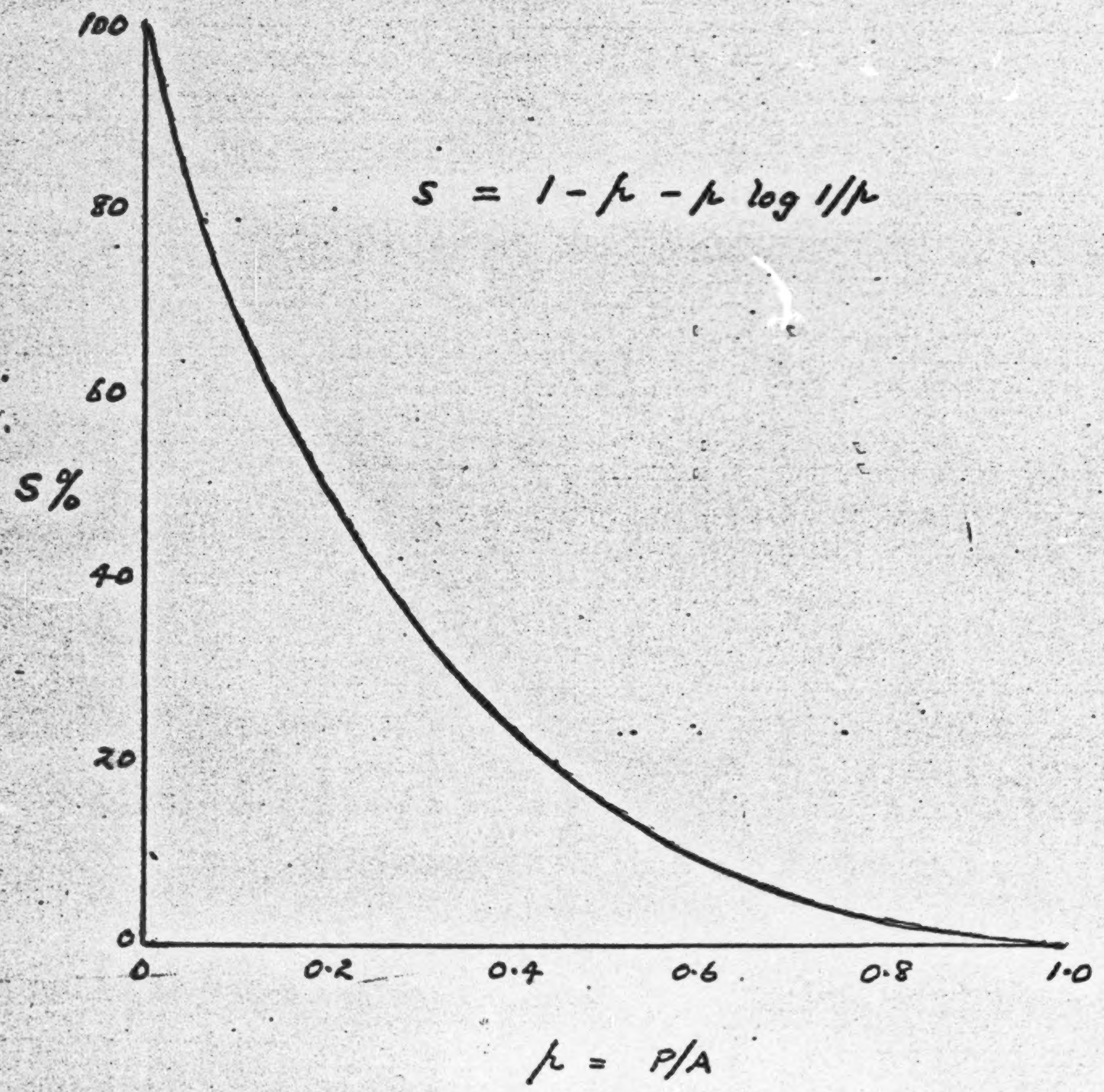
ASLIB PROCEEDINGS: EAST, WEYMAN
APRIL 69.

N = 248 [372]

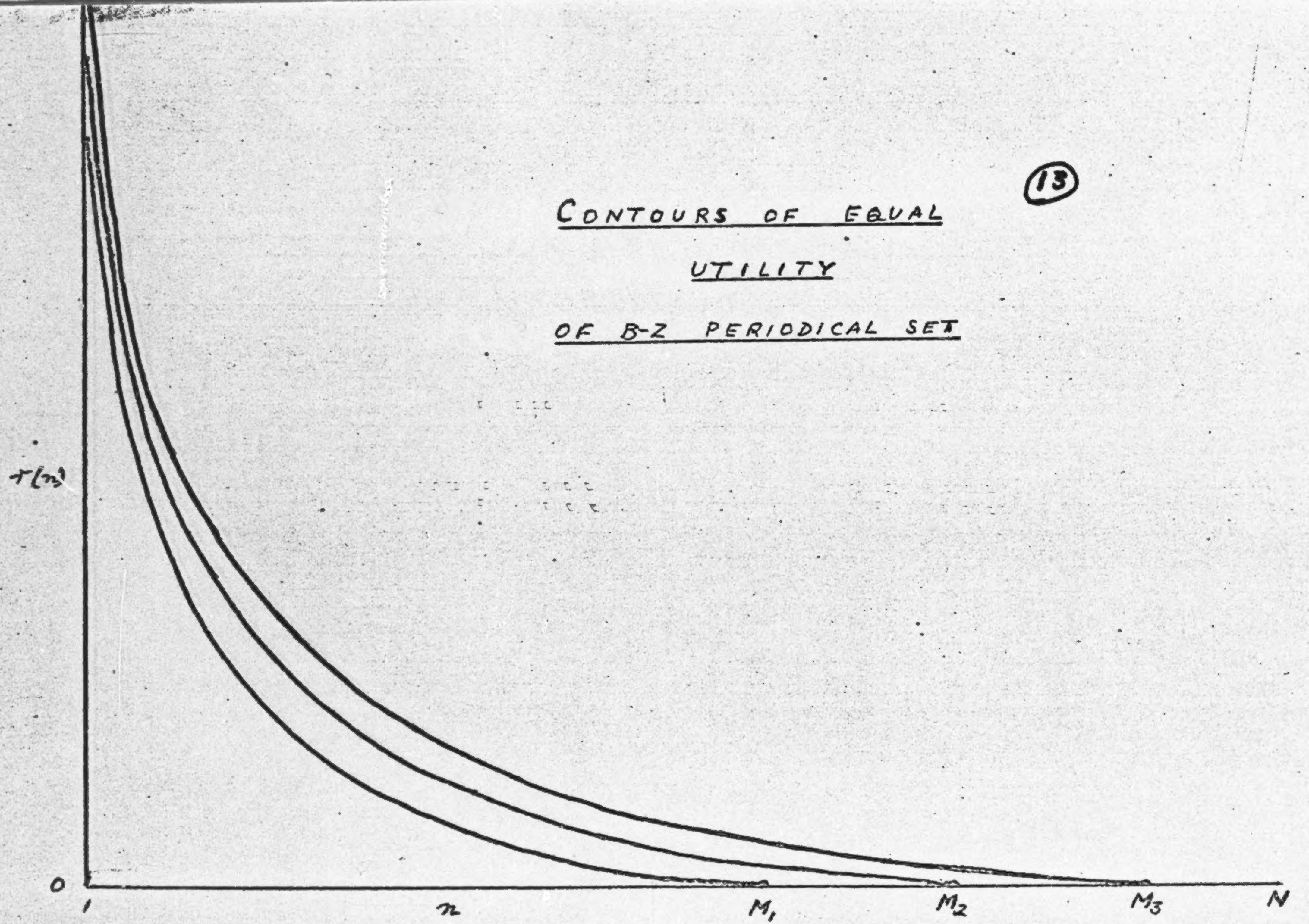
PHOTOCOPIES:

SAVING AT OPTIMUM CUT-OFF

⑫



$$S = 1 - p - p \log 1/p$$



⑬

CONTOURS OF EQUAL UTILITY OF B-Z PERIODICAL SET

STORAGE POLICIES FOR INFORMATION SYSTEMS

by

Ferdinand F. Leimkuhler°

ABSTRACT

The cost of an item of information in a storage system is defined as the sum of an initial cost, a time dependent cost, and a usage cost. Item usage is assumed to follow a simple exponential obsolescence pattern. A decision rule for the economic holding time of an item in storage is derived from the model and based on a policy of minimizing the average cost of usage. Some properties and implications of the rule and policy are discussed. This model is developed in the context of large research libraries but should be applicable to other kinds of information systems.

°The writer is a professor of industrial engineering at Purdue University, and a visiting professor of librarianship at the University of California, Berkeley, for 1968-69. This work is supported in part by NSF grant GN-759 at Purdue and the Ford Foundation Research Project 68-267 at the University of California. This paper is prepared for the Joint National Meeting of the American Astronautical Society and the Operations Research Society at Denver, June 1969, and for the Library Planning Conference at the University of Lancaster, England, July, 1969.

Introduction

It has been pointed out by C. W. Churchman [1968] that the more important issue in current efforts to develop automated information systems is not how to merely computerize existing clerical practices, but how to use this new technology to enlarge our concept of information systems so as to include the user as a more integral and active component. If libraries, accounting systems, or other kinds of information systems are to be made more responsive to the needs of the persons using them, then the appropriate level of planning and control must be broad enough to include the user as an effective participant. Churchman goes on to point out that even a well-designed user-library system would be a subsystem in a larger environment and would run the risk of suboptimizing its policies relative to the larger system.

N. R. Baker [1967] has suggested one way to formalize the notion of system expansion by defining a service system as composed of three active components: the service agency proper, its users, and its funders. The funder component is the final arbiter of system performance. This model permits a closed-loop analysis of the interactions among the three groups, which Baker used to show that the "servicers" can expect to become increasingly more constrained in their decision-making unless they can make satisfactory decisions before users exercise their influence and before funders exercise their powers of control to force the decisions. The service agency must convince the others that learning, not influence and control, is the dominant factor and the most productive approach; and they must demonstrate the fact that they are exploiting fully the political, economic, and technical resources which are available to them.

Libraries are among the world's largest information systems. They provide a rich history of operational experience for the student of general information systems and a large working environment in which to test new design concepts. Although conventional libraries are essentially manual systems for the handling of mechanically-stored information in book form, many of their operating characteristics are readily transferable to more sophisticated systems using computers and microform storage devices. This is most apparent in the library operations research studies of recent years, and the seminal work in this field is reported in the recent monograph by P. M. Morse [1968]. Library operations research studies have concentrated on the problems of storing and using library materials, while library and information scientists have focused on problems of organizing and retrieving these materials according to their intellectual content. The latter problems seem to constitute a more difficult long-run research field, since the introduction of the newer methods of information storage preclude direct user access and require newer methods of obtaining remote intellectual access to the file.

Much of the operational analysis of libraries is related directly to the problem of library size, and the use of such options as depositories, interlibrary loans, blanket orders, duplication, and compact storage, as means of optimizing library size relative to the observed usage of the library. Usually, the library under study is thought of as a member of a larger information network which permits local suboptimization without precluding the possibility of the user going elsewhere for information. A good prototype example of this kind of approach is the model proposed by P. F. Cole [1962], and refined by M. K. Buckland and I. Woodburn [1968], by which it is shown that a 2,000 volume petroleum library can expect to satisfy the greatest number of user requests by subscribing to approximately 190 different journals or serials and holding them for about eleven years. Variations on this theme of "optimal library size" are seen in the study of depositories by Morse [1968] and W. C. Lister [1967] and the study of interlibrary loan by G. Williams [1968]. A more sophisticated approach is the fully stochastic model of H. M. Gurk and J. Minker [1968] which studies the effect of retention policies on the size of a data base for a computer utility.

The size of a library or data base seems to be the most important measure of its worth apart from its usage, since it suggests comprehensiveness or completeness of knowledge. This has long been the traditional measure of stature in library circles. The two important determining factors of size are the breadth of acquisition and the length of retention. These are also important factors in determining usage along with the ease of access. While some models have been developed which concentrate on library breadth, (see Leimkuhler [1967, 1968]), the problem of retention time has been given the greatest attention. The storage cost model and storage policy developed below is intended to reveal some of the essential economic characteristics of information storage systems in an elementary way by developing a decision rule for the holding time which is both practicable and near-optimal.

Cost of Storing a Single Item

Recent studies of the cost of operating library-type information systems, such as the work of Williams [1968], and R. Shisko [1968], suggest the following cost model for information storage systems:

$$K(t) = k_1 + k_2 t + k_3 u(t) \qquad (1)$$

Here $K(t)$ represents the total cost of holding one item for a period of t years; k_1 is the initial cost of acquiring the item; $k_2 t$ is the holding cost which is linearly related to the retention period; and $k_3 u(t)$ is the usage cost which is proportional to the number of uses made of the item during the period t. This model is consistent with those used by Lister [1967] and Buckland [1968], although their models included more terms in order to recognize other control variables. Equation (1) could be discounted in order to obtain its equivalent present value as was done in the study by Williams. Equation (1) is not supposed to represent the ordinary way in which the costs of libraries or other types of information systems are reported for either budgetary or cost control purposes. Rather, it is intended to express storage cost as a function of time and usage in the simplest possible manner. There is no theoretical reason, for example, for not including user costs in the parameters along with the direct and indirect costs of the storage system proper.

In his study of book use models, A. K. Jain [1967] described several models which express book usage as a function of age. In all of these models the cumulative use, $u(t)$, increases monotonically with t, while $u'(t)$ decreases. The simplest of these models is the exponential case, that is:

$$u'(t) = re^{-bt} \qquad (2)$$

$$u(t) = (r/b)(1 - e^{-bt}) \tag{3}$$

where r is a scale parameter associated with the instantaneous initial usage level and b denotes the instantaneous obsolescence rate. The ratio (r/b) is the limit of $u(t)$ as t approaches infinity and therefore a measure of the lifetime usage of the item. Based on an extensive study of the M. I. T. Libraries, Morse [1968] proposed a usage model similar to that of equation (3) but including a constant or residual use term which is independent of age, that is, the usage rate drops exponentially to a residual level. He showed that this model results from a simple Markov process for the change in usage from year to year.

By substituting equation (3) into equation (1), the total, marginal, and average costs as a function of holding time are obtained respectively, as follows:

$$K(t) = k_1 + k_2 t + k_3 (r/b)(1 - e^{-bt}) \tag{4}$$

$$K'(t) = k_2 + k_3 r e^{-bt} \tag{5}$$

$$\overline{K}(t) = (k_1/t) + k_2 + k_3 (r/bt)(1 - e^{-bt}) \tag{6}$$

Both the marginal cost and average cost of retention time diminish to the level k_2 as the holding period increases, and the total cost becomes increasingly linear with time.

Cost of Providing Uses of an Item

A more interesting and useful cost relationship is obtained by expressing the total cost as a function of the cumulative usage during the retention period. By inverting equation (3), one obtains the time required to provide the first u uses of an item in storage, that is:

$$t(u) = \ln(1 - bu/r)^{-1/b} = (-1/b)\ln(1 - bu/r) \qquad (7)$$

By substituting equation (7) into equation (1), the total cost for providing the first u uses is defined as follows:

$$K(u) = k_1 - (k_2/b)\ln(1 - bu/r) + k_3u \qquad (8)$$

The marginal cost for providing the u^{th} service is approximately equal to

$$K'(u) = k_3 + k_2/b(1 - bu/r) \qquad (9)$$

where it is assumed that the derivative of $K(u)$ approximates the finite difference, $K(u) - K(u-1)$. The average cost of providing the first u uses of an item is defined by the equation:

$$\overline{K}(u) = (k_1/u) - (k_2/bu)\ln(1 - bu/r) + k_3 \qquad (10)$$

While both the total cost and marginal cost of usage increase monotonically and quite rapidly with increased usage, the average cost decreases at first and then increases with usage.

The implications of equations (8), (9), and (10) can be more readily seen if they are expressed in terms of a relative measure of usage, x , which is the ratio of the cumulative usage over the lifetime usage, that is:

$$x = bu/r \tag{11}$$

It is convenient also to define the parameters K_2 and K_3 as follows:

$$K_2 = k_2/b \tag{12}$$

$$K_3 = rk_3/b \tag{13}$$

where K_3 can be interpreted as the total lifetime usage cost of an item, and K_2 as the holding cost for a relaxation interval, $1/b$. By using these definitions, the equations for the total, marginal, and average cost of usage become:

$$K(x) = k_1 - K_2\ln(1 - x) + K_3x \tag{14}$$

$$K'(x) = K_3 + K_2/(1 - x) \tag{15}$$

$$\overline{K}(x) = (k_1/x) - (K_2/x)\ln(1 - x) + K_3 \tag{16}$$

These relationships are plotted in Figure 1 to show their general shape and properties. The plotted values are based on the arbitrary assumption that k_1, k_2, and K_3 are of equal magnitude.

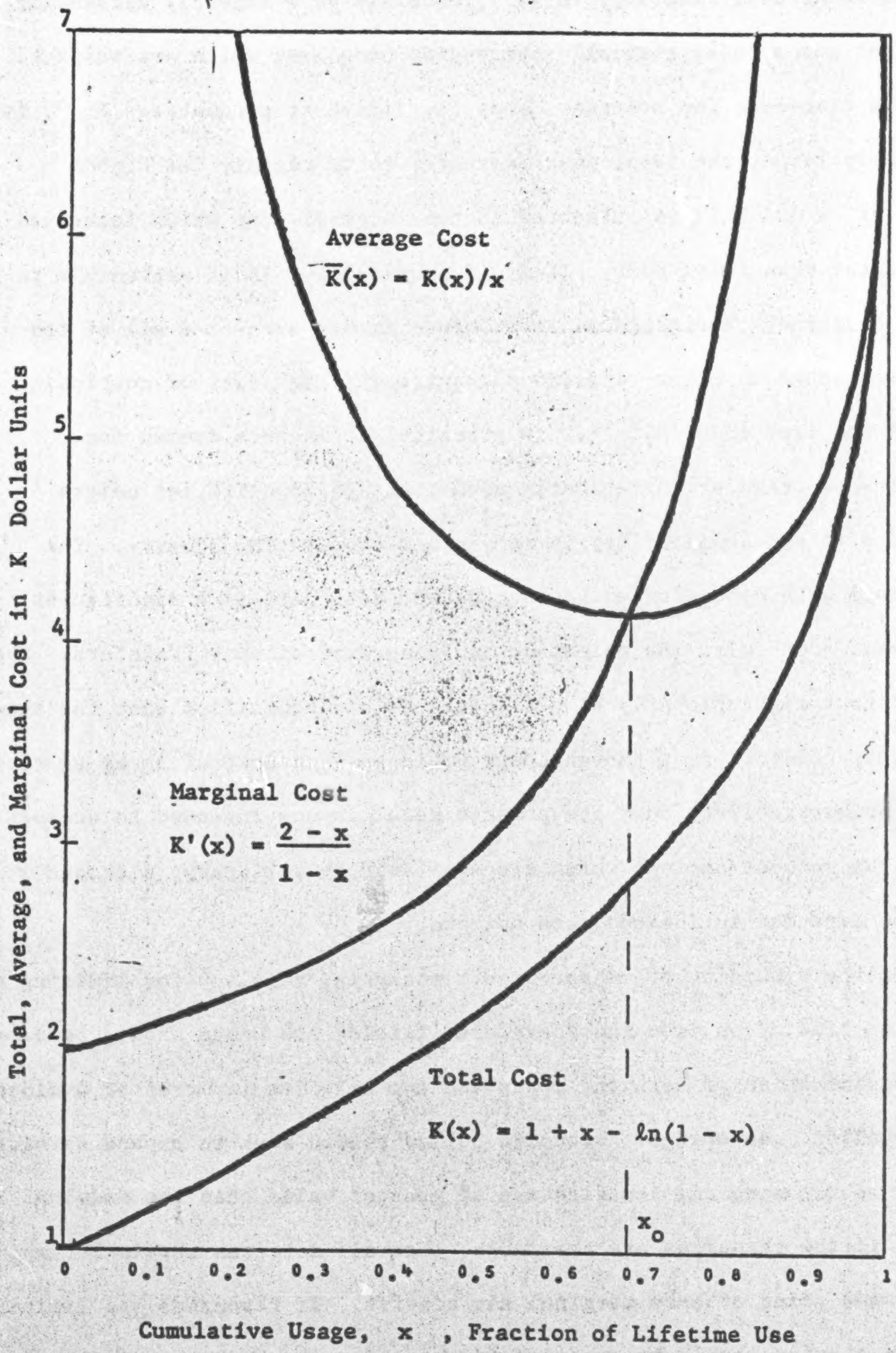


Figure 1--Information Storage Costs of One Item as a Function of its Usage, When Parameters k_1 , K_2 , and K_3 Are Equal to K .

Storage Policies for a Single Item

The total cost function, $K(x)$, consists of a linearly increasing component and a logarithmically increasing component which are weighted with the time-cost for storage. When the time-cost parameter, K_2 , is relatively large, the total cost increases quite rapidly for higher values of x . This is reflected in the marginal cost which increases much faster than total cost. If it is permissible, it is reasonable to expect a library to discard an item before it has exhausted all of its potential usage in order to avoid the extremely high cost of continuing to hold the item indefinitely. In practice, it is more common for libraries to transfer infrequently used items to depositories unless assured of their availability in some other cooperating library. The experience with depositories has suggested that there is a significant cost associated with the selection and recording of such transfers. Much of this cost might properly be considered as an acquisition cost for the depository collection, although there would be some cost of changing records in the primary collection. The present model is not intended to account for all of the various options which are available to a library, although it could be expanded to include such options.

From the viewpoint of microeconomic analysis, a policy for limiting the retention time of an item and therefore limiting its usage should be based on a consideration of both the costs and the benefits incurred or avoided by the policy. An optimal economic policy should seek to expand service as long as the marginal benefits are of greater value than the marginal costs. If the resources are available, then all services should be expanded to the same point of zero marginal net benefit. If resources are limited, then the service should be expanded to the point where the marginal net benefit is the same for all costs, since, otherwise, the costs could be reallocated so as to increase the total net benefit. In order to apply these optimality principles directly one needs to evaluate the benefits derived from item usage in a manner which is directly comparable to the cost measurements. However, the direct measurement of the economic value of the benefits of information retrieval is an extremely difficult, if not impossible, task, and indirect methods are the only recourse.

An alternative approach to the establishment of storage policies is to choose that retention period which minimizes the average cost of usage. In addition to the practical advantage of being based on the direct measurement of costs only, this policy has economic attributes which recommend it as a near-optimal solution with regard to user benefits, also. There is good reason to suppose that the marginal and average benefits from item usage are relatively constant from the standpoint of anticipating such benefits for the purpose of establishing a policy. Furthermore, average benefits should be at least as great as the average cost in order that the entire venture is not unprofitable. By holding an item to the point where average cost is minimized, there it is an assurance that at least a break-even in the cost-benefit relationship has been achieved. This is a relatively conservative approach to the problem which is not at all unreasonable when there is almost complete ignorance about the relative worth of the benefits derived from item usage.

There is a well-established economic thesis which holds that the long-term tendency in competitive production is for the producers to be driven to the point of zero net profit, that is, where average cost equals average revenue. While the situation in information storage is not directly analogous, it seems to be quite similar in that there are usually alternative information sources available to the user, and these alternatives will be exercised as long as they can do so at less cost. The competitive interaction of users and suppliers should tend to match benefits with costs. Another argument in favor of a minimum average cost policy is that it is a highly practical operational policy for a production or service subsystem to follow, since it motivates local efficiency and technical innovation. For example, it provides a viable guide to the management of a factory in meeting its production goals at minimum cost and for reporting to higher management the factory data they need to establish goals. Standard cost accounding procedures develop average cost figures which become a measure of factory performance.

Minimization of the Average Cost of Usage

A storage policy based on the minimization of the average cost of usage is relatively easy to implement on the basis of cost information alone. Since the average cost achieves a minimum value when it is equal to the marginal cost, a decision rule can be easily obtained by equating equations (15) and (16) and solving for x as follows:

$$\text{Min } \overline{K}(x) \;\Rightarrow\; k_1/K_2 \;=\; \ln(1 - x) + x/(1 - x) \qquad (17)$$

This decision rule is evaluated in Table 1 where the relationship is shown between the parametric ratio (k_1/K_2) and the value of x which minimizes average cost. By referring to equation (3), it is possible to translate this decision rule into the holding times which minimize average cost as follows:

$$\text{Min } \overline{K}(x) \;\Rightarrow\; k_1/K_2 \;=\; e^{bt} - 1 - bt \qquad (18)$$

where b is the obsolescence rate and bt expresses holding time in the number of relaxation intervals. By expressing holding time this way, it is possible to demonstrate the effect of the decision rule on holding time using equation (7). This is done in Table 1.

An approximate version of the decision rule can be obtained by expanding the exponential term in equation (18) and ignoring all but the first three terms in the expansion. This leads to the simpler rule:

$$\text{Min } \overline{K}(x) \;\Rightarrow\; t_h \;=\; \sqrt{2k_1/bk_2} \qquad (19)$$

where t_h denotes a holding time which effects an approximate minimization of the average cost of usage. This version of the decision rule has some intuitive appeal because of its similarity to the economic lot-size formula of inventory theory. The control parameter t_h can be called the "economic holding time" for an information system. It can be seen in Table 1 how t_h tends to overestimate the time required to minimize average cost especially at unusually large values of the ratio (k_1/K_2). Equation (19) implies that the economic holding time will change as the square root of changes in the cost of acquisition and storage or changes in the obsolescence rate. As the cost of acquisition, k_1, decreases the holding time will decrease, and as the storage cost, k_2, decreases, the holding time is increased. An increase in the obsolescence rate will decrease the holding time and will decrease the total usage obtained from the item since, in equation (17), an increase in b decreases the parameter K_2, which decreases both the ratio (k_1/K_2) and the minimizing value of relative usage, x.

Table 1. Values of Relative Usage, Holding Times, and Costs which Minimize the Average Cost of Usage for an Item

Fraction of Total Lifetime Usage x	Holding Time (relaxation intervals) bt	Ratio of Cost Parameters k_1/K_2	Economic Holding Time as Computed from Equation (19) $bt_h = \sqrt{2k_1/K_2}$
0.1	0.11	0.01	0.11
0.2	0.22	0.03	0.23
0.3	0.36	0.07	0.38
0.4	0.51	0.16	0.56
0.5	0.69	0.31	0.78
0.6	0.92	0.58	1.08
0.65	1.05	0.81	1.27
0.70	1.20	1.13	1.50
0.75	1.39	1.61	1.80
0.80	1.61	2.39	2.19
0.85	1.90	3.77	2.75
0.90	2.30	6.70	3.67
0.95	3.00	16.00	5.66
0.99	4.61	94.40	9.72

Concluding Remarks

It is interesting to note that the decision rule establishes the holding time independently of the usage parameter, k_3 . In fact, if only acquisition and time dependent costs are considered, the holding time would be the same. The interesting point is that it is reasonable to argue that almost all of the costs of operating a library can be allocated between these two cost categories, since most of the labor cost in libraries is expended for professional or semi-professional personnel who in many ways represent as much of a system investment as do the purchase price of the materials. Almost all categories of library cost correlate closely with the size of the collection and/or the acquisition rate of new materials. Even the acquisition costs are correlated closely with size, because of the steady exponential growth patterns which are characteristic of large libraries. Some, but certainly a small part, of direct library expense does vary directly with usage, as in the operation of reserve book rooms where items circulate with a very high frequency. If it is valid to consider storage system costs as being represented by the parameters k_1 and k_2 , only, then it would seem worthwhile to consider the cost parameter k_3 as being representative of the cost to the user in obtaining information from the system. Equations (8) or (14) would then represent the combined total cost to both the patron and the storage system for providing uses from an item, and the decision rule would determine the holding time which minimizes the combined average cost per use.

An alternative interpretation would be to consider the user cost, k_3 , as a monetary estimate of the benefit to the user for his effort in using the item, under the assumption that the user chooses among alternative sources in such a way as to eventually equate benefits with costs on the average. Under this interpretation, it would be appropriate to exclude k_3 from the equations and compare the marginal and average costs with the parameter k_3 . The decision rule for holding time would be the same, but if it yielded a minimum average cost which is less than k_3 , that is, if the benefit/cost ratio is greater than one, there would be an indication that the system is not operating at an optimal level of service and should be expanded beyond the point of minimum average cost. This interpretation seems to be in line with the arguments of R. L. Meier [1961], who found that university libraries were operating in a range where the average cost per unit of service was the same as the average cost to a student patron. However, he also found that the average cost to a faculty patron was much higher than the library cost figure, and that this situation was discouraging faculty patronage. He concluded that libraries should expand their services and absorb more of the faculty usage cost so as to obtain a net increase in total benefit to the university.

The model and policy advocated here and its accompanying speculation have not been fully tested by either analytic methods or by comparison with empirical data. They are offered as an alternative approach to the rational management of large library-type information storage systems. Constructive criticisms are invited. Although this model has been developed in the context of large research libraries, there is no evident reason why its line of argument would not be directly applicable to other types of information storage systems.

REFERENCES

Baker, N. R. [1967] "Toward an Analytic Framework for Service Organizations," Paper, 14th International Meeting, Institute of Management Sciences, Mexico City.

Buckland, M. K., and Woodburn, I. [1968] "Some Implications for Library Management of Scattering and Obsolescence," Report, University of Lancaster Library, England.

Churchman, C. W. [1968] "Management Information Systems," Chapter 7, in The Systems Approach, Dell Publishing Co., New York.

Cole, P. F. [1962] "A New Look at Reference Scattering," J. of Documentation, Vol. 18, No. 2, pp. 58-64.

Gurk, H. M., and Minker, J. [1968] Storage Requirements for Information Handling Centers, Tech. Report No. 68-61, University of Maryland Computer Science Center.

Jain, A. K. [1967] "A Statistical Study of Book Use," Thesis, Purdue University, and U. S. Clearinghouse Report No. PB 176525, Springfield, Virginia.

Leimkuhler, F. F. [1967] "The Bradford Distribution," J. of Documentation, Vol. 23, No. 3, pp. 197-207.

Leimkuhler, F. F. [1968] "A Literature Search and File Organization Model," American Documentation, Vol. 19, No. 2, pp. 131-136.

Lister, W. C. [1967] "Least Cost Decision Rules for the Selection of Library Materials for Compact Storage," Thesis, Purdue University, and Clearinghouse Report No. 174441, Springfield, Virginia.

Meier, R. L. [1961] "Efficiency Criteria for the Operation of Large Libraries," Library Quarterly, Vol. 31, pp. 215-234.

Morse, P. M. [1968] Library Effectiveness: A Systems Approach, M.I.T. Press.

Morse, P. M., and Elston, C. [1969] "A Probabilistic Model for Obsolescence," Operations Research, Vol. 17, No. 1, pp. 36-47.

Shisko, R. [1968] "The Application of Systematic Analysis to the M.I.T. Libraries," Essay, Department of Political Science, Massachusetts Institute of Technology.

Williams, G. [1968] Library Cost Models: Owning Versus Borrowing Serial Publications, Report, Center for Research Libraries, Chicago.

A MATHEMATICAL MODEL OF A HIERARCHICAL LIBRARY[x] SYSTEM

I. Woodburn
University of Lancaster

Summary

This report describes a mathematical model for a hierarchical library system in which books, journals, etc. are stored at different levels.

It illustrates by means of a simplified example how such a model can be used to design an efficient library system.

The material is presented in the context of a library system for the universities.

---

[x] The work described in this paper forms part of the 'Systems analysis of a university library' project at the University of Lancaster Library and was supported by a grant from the Office of Scientific and Technical Information.

CONTENTS

Page

1. Introduction 3

2. The basic model 4

   2.1. Application of the model 6

   2.2. Operating costs 8

   2.3. Access times 9

   2.4. User costs 10

3. Extensions of the model 13

   3.1. Elasticity of demand 13

4. Conclusions 14

5. References 14

Introduction

The library user at a university obtains information from books, etc. stored at several points in a hierarchical system. In the majority of cases this system has four levels:

(1) a personal collection

(2) a departmental library

(3) a university library

(4) other university libraries and national libraries

Because material is more accessible at the lower levels (i. e. personal collections and departmental libraries) it is reasonable to suppose that, from the point of view of the individual, the best system at a given cost is one in which the most-used material belongs to his personal collection and so on by degrees until the least-used material is held at national level.

Although a personal library is established for the benefit of one individual, a departmental library must cater for demands from all its members. In practice it must cater for the sum of the residual demands from its members since some of their demands will have been met from their own collections. Similarly a university library must cater for the sum of the residual demands from all members of all its departments and a national library for the sum of residual demands from all universities.

This principle, the principle of residual demand, has been used to establish a simple mathematical model for a hierarchical library system. The model will be described in mathematical terms and its use will be illustrated by means of a simplified but realistic example.

2. THE BASIC MODEL

Consider a library system which has four levels of storage:-

(1) Personal libraries
(2) Departmental libraries
(3) University libraries
(4) a National library

Let M be the total number of titles to be stored in the system, and let N be the total number of users of the system.

Let r^i_{jkl} be the expected number of demands in a specified period of time (e.g. one year) for the ith title from the jth user belonging to the kth department of the lth university.

Considering each storage level in turn:-

(1) Personal library: we assume that personal libraries hold all titles i for which

$$r^i_{jkl} \geqslant r^x$$

NB r^x is the marginal level of demand below which the title is not held (for simplicity it is the same for all titles)

(2) Departmental library: let $r^i_{.kl}$ be the expected number of demands for the ith title from all members of the kth department of the lth university.

Then applying the principle of residual demand we have

$$r^i_{.kl} = \sum_j r^i_{jkl}$$

where $r^i_{jkl} = o$ if $r^i_{jkl} \geqslant r^x$ at the personal library level.

We assume that departmental libraries hold all titles i for which

$$r^i_{.kl} \geqslant r^{xx}$$

NB r^{xx} is the marginal level of demand below which the title is not held and it is the same for all titles.

(3) University library:

Let $r^i_{..l}$ be the expected number of demands for the ith title from all members of all departments of the lth university.

Then applying the principle of residual demand we have

$$r^{i}_{..l} = \sum_{k} r^{i}_{.kl}$$

where $r^{i}_{.kl} = o$ if $r^{i}_{.kl} \geqslant r^{xx}$ at the departmental library level.

We assume that University libraries hold all titles i for which

$$r^{i}_{..l} \geqslant r^{xxx}$$

NB r^{xxx} is the marginal level of demand below which the title is not held and it is the same for all titles.

(4) National library: at this stage we have a choice, we can insist that the national library holds all titles or we can define r^{xxxx} to be the marginal level of demand below which the title is not held and apply the principle of residual demand to discover which titles should not be stored at the national level. The first assumption seems more appropriate and as the costs of duplication are small we assume that the national library holds all M titles.

At this stage it should be evident that the marginal levels of demand r^{x}, r^{xx} and r^{xxx} constitute a storage policy for the system - if we change their values then we change the pattern of storage and also the way in which demands are satisfied. For any particular policy we can calculate the number of titles stored at each level and also the total number of demands satisfied from each level, and the total cost of providing the service can be calculated. Also if we define t^{x}, t^{xx}, t^{xxx} and t^{xxxx} to be the average times of access to material stored at personal, departmental, university and national levels respectively, then we can calculate the overall average time of access to the system.

One simplification that has been made at the outset is that material can only be transferred between libraries at distinct levels in the system e.g. inter University library loans are treated as loans from a national library as far as cost and access times are concerned.

If we wish to compare two storage policies then we select two sets of values for the marginal levels of demand and simulate the operation of each system for a fixed pattern of expected demands r^{i}_{jkl}. The total cost and overall average access times for each system can then be compared.

The applicability of the model will be illustrated by means of a simplified but realistic example which highlights the information that is required before systems of this kind can be viewed objectively.

2.1. Application of the model

There are about 50 Universities with an average of 30 departments per University and a grand total of 250, 000 staff and students. It is most unlikely that the expected demand for each title from each of the 250, 000 users can be measured and an alternative approach is necessary. Personal libraries are excluded from the illustrative example so that the basic quantities required are the $r^i_{.kl}$ i.e. the summed residual demands for the ith title from all members of the kth department of the lth university.

Even so there are about 7 million titles held by the British Museum library at the present time and it is most unlikely that the expected demands for each of these titles can ever be measured. Therefore in the example the number of titles that has been considered is restricted to what might be called "the most used scientific journals in the early 1950's".

The data that has been used to represent the quantities $r^i_{.kl}$ has been taken from the book by C.H. Brown Scientific Serials (ACRL monograph, 16) Chicago, ACRL. 1956. In the course of Brown's study he produces a consolidated alphabetical list of the 612 most frequently cited scientific serials classified according to the source journals in eight subject fields. The subject fields together with the total number of journals cited and the total number of citations are listed in Table 1.

Table 1

	No. journals cited	No. citations
(1) Mathematics	179	3348
(2) Physics	320	9596
(3) Chemistry	275	10518
(4) Geology	490	2913
(5) Physiology	299	5984
(6) Botany,	376	4995
(7) Zoology	663	2775
(8) Entomology	350	2326
TOTAL	2952	42455

The figure of 2925 journals cited includes some duplication but the necessary data is not recorded. However it is known that there were a total of 828 journals cited five or more times and that these figures refer to the 612 distinct titles in the consolidated list. It has been assumed in the exercise that there is no overlap in the less cited journals i. e. they are classified into one subject heading only. This implies that 2925 - 828 + 612 = 2736 distinct journals were cited a total of 42455 times.

The degree of overlap of the 2736 cited journals can be measured by the number cited in 1, 2, 3 etc. subject fields. The complete distribution is contained in Table 2.

Table 2

No. subject fields	No. journals cited
1	2617
2	67
3	34
4	6
5	4
6	4
7	2
8	2
	2736

The apparent lack of overlap in the basic data helps to explain the large part played by departmental libraries in the ensuing analysis.

The basic data for the simulation exercise is the array of numbers listed by Brown on pages 143-154 of his book, and the first ten rows of the array are reproduced in Table 3 for the purpose of illustration.

Table 3

	Maths.	Physics	Chem.	Geol.	Physiol.	Botany	Zool.	Entom.	TOTAL
Acad. Med. Belg. Bull.	.	.	.	.	10	.	.	.	10
Acad. Belg. Cl. Sci. Bull	12	.	.	.	.	.	.	.	12
Acad. Sci. Paris, C.R.	125	213	239	16	19	173	43	16	844
Acad. Nat. Sci. Philad., Proc.	.	.	.	.	.	.	8	.	8
Acad. Lincei, Mem.	.	.	8	.	.	.	.	.	8
Acad. Lincei, R.	10	.	.	.	.	.	.	.	10
Acoust. Soc. Amer., J.	.	16	.	.	.	.	.	.	16
Acta chem. scand.	.	.	212	.	.	.	.	.	212
Acta cryst., Camb.	.	32	14	9	.	.	.	.	55
Acta Hort. berg.	.	.	.	.	.	11	.	.	11

In this exercise the array represents the quantities $r^{i}_{.kl}$ with the index i running over the 612 journals (or more accurately, the 2736 journals). The subject fields represent the departments i.e. k = 1, 2, 3, ..., 8.

For the sake of extreme simplicity it has been assumed that the library system mus cater for 50 identical universities, each with 8 departments, and with identical demands for each of 2736 titles i.e.

$$r^{i}_{.kl} = \text{constant for all } l$$

Since the degree of overlap in expected demands is more important to the design of an efficient storage system, this simplification is not critical.

The time period is considered to be one year so that 42455 demands are made in one year for 2736 titles at each of 50 Universities.

Another simplification has been made in defining alternative storage policies for the system, namely that the marginal levels of demand are equal at each level of storage i.e. $r^{xx} = r^{xxx}$. In order to provide a wide range of results the levels that have been investigated are 5000, 2000, 1000, 500, 200, 100, 50, 20, 10 and 5 demands per annum.

Table 4 contains the number of titles stored and the number of requests satisfied from each storage level for each of these marginal levels of demand. The results apply to each of the 50 universities.

Table 4

Storage level

Marginal level of demand	Departmental Titles	Requests	University Titles	Requests	National Titles	Requests
5000	NIL	NIL	NIL	NIL	2736	42455
2000	2	5606	NIL	NIL	2736	36849
1000	3	6762	NIL	NIL	2736	35693
500	7	9876	4	3105	2736	29474
200	38	18551	10	2940	2736	20964
100	68	22667	16	2329	2736	17459
50	135	27409	30	2169	2736	12877
20	308	32544	43	1186	2736	8725
10	568	36022	9	127	2736	6306
5	612	37834	NIL	NIL	2736	4621

2. 2. Operating costs

We assume that each demand generates one issue and return and that the unit cost of the pair of transactions is 1/- from whatever storage level that the demand is satisfied.

In addition a transfer charge is included for all demands satisfied from the national level, this covers postage and packing and a handling charge. It is not a fixed unit cost since economies of scale are important. A curve has been drawn for which the unit cost is 15/- when there are 150, 000 loans per annum, 10/- when there are 460, 000 loans per annum, and 5/- when there are more than 2, 000, 000 loans per annum.

At this stage another simplification has been achieved by assuming that a single copy of each journal is sufficient to provide an adequate standard of service so far as the "immediate" availability of material is concerned.

These unit costs have been used to derive the following table of operating costs for the system as a whole.

Table 5

Marginal level of demand	Purchase	Costs (£ per annum) Issue	Transfer	TOTAL
5000	41, 040	106, 138	530, 688	677, 866
2000	42, 540	106, 138	460, 612	609, 290
1000	43, 290	106, 138	446, 162	595, 590
500	49, 290	106, 138	386, 846	542, 274
200	77, 040	106, 138	327, 562	510, 740
100	104, 040	106, 138	290, 983	501, 161
50	164, 790	106, 138	265, 588	536, 516
20	304, 290	106, 138	223, 578	634, 006
10	473, 790	106, 138	189, 180	769, 108
5	500, 040	106, 138	155, 959	762, 137

The minimum operating cost is associated with the marginal level of demand of 100 which, on referring back to Table 4, is equivalent to a 40% level of loans from the national library (i. e. 40% of total loans are satisfied from the national library level). In fact costs do not vary greatly over the range 30-50% national library loans.

2. 3. Access times

So far the only costs that have been considered are the costs of providing the service. We now go on to consider the costs that are incurred by the user and we begin by defining average times of access. These times are associated with the time it takes to obtain a book or a journal after the need for it has been recognised.

Times assumed are:-

0.1 hours from a departmental library

0.5 hours from a University library

12 hours from a national library

(this last estimate represents $1\frac{1}{2}$ working days).

The overall average access times for the total number of demands are calculated from the results listed in Table 4 to be:

Table 6

Marginal level of demand	Overall average access times (hours)	Total annual operating costs (£)
5000	12	677, 866
2000	10.4	609, 290
1000	10.0	595, 590
500	8.4	542, 274
200	6.0	510, 740
100	5.0	501, 161
50	3.7	536, 516
20	2.5	634, 006
10	1.9	769, 108
5	1.4	762, 137

2.4. User costs

At this stage we have attempted to cost users time in money terms so that an overall assessment of different storage policies can be made. However, because of the difficulties in estimating such costs it has been necessary to propose a number of alternative costings and examine the consequences of these. The costs that were selected are:-

40/- per hour

4/- per hour

(additional unit costs are considered in a later section of this report)

In order not to penalise the national library unfairly, the access time for costing purposes was reduced to 1 hour (i. e. two visits to the University Library) since it is reasonable to assume that the remainder of the 12 hours wait would be used profitably.

The total access times for all demands from each university based on this revised figure, the annual user cost, the annual operating cost and the grant total cost are listed in Tables 7 and 8 for unit user costs of 40/- per hour and 4/- per hour respectively.

Table 7

User cost 40/- per hour

Marginal level of demand	Total access time (hours) each univ.	Costs (£ per annum) user	operating	grand total
5000	42, 455	4, 245, 500	677, 866	4, 923, 366
2000	37, 409.6	3, 740, 960	609, 290	4, 350, 250
1000	36, 369.2	3, 636, 920	595, 590	4, 232, 510
500	32, 014.1	3, 201, 410	542, 274	3, 743, 684
200	24, 289.1	2, 428, 910	510, 740	2, 939, 650
100	20, 890.2	2, 089, 020	501, 161	2, 590, 181
50	16, 702.4	1, 670, 240	536, 516	2, 206, 756
20	12, 572.4	1, 257, 240	634, 006	1, 891, 246
10	9, 971.7	997, 170	769, 108	1, 766, 278
5	8, 404.4	840, 440	762, 137	1, 602, 577 *

Table 8

User cost 4/- per hour ~~Costs (£ per annum)~~

Marginal level of demand	Costs (£ per annum) User	Operating	Grand total
5000	424, 550	677, 866	1, 102, 416
2000	374, 096	609, 290	983, 386
1000	363, 692	595, 590	959, 282
500	320, 141	542, 274	862, 415
200	242, 891	510, 740	753, 631
100	208, 902	501, 161	710, 063
50	167, 024	536, 516	703, 540 *
20	125, 724	634, 006	759, 730
10	99, 717	769, 108	868, 825
5	84, 044	762, 137	846, 181

The effect of including user costs is to reduce the part played by the national library. If the user cost is 4/- per hour then about 30% of loans should be satisfied from the national level, but if it is 40/- per hour then less then 10% of loans should be satisfied from this level.

3. EXTENSIONS OF THE MODEL

3.1. Elasticity of Demand

One feature that has been excluded concerns the reduction in demand which occurs when material is not readily accessible. For example, an individual who has good departmental library facilities will think twice about visiting the University library to obtain a book or a journal.

It is quite easy to introduce a factor for this elasticity in demand for each storage level which would have the effect of reducing the summed residual demands to be catered for at higher storage levels. It would also be possible to conclude a total cost analysis of alternative storage policies provided that the "lost" information could be valued in money terms.

4. CONCLUSIONS

The intention of this report has been to describe a simple model of a hierarchical library system which can be used as a basis for further discussion.

It is stressed that the results which have been obtained are fictitious because accurate data is not available at the present time to permit a genuine enquiry into storage policies. The example has only served to highlight those areas in which data is needed before a proper study can be undertaken.

The most pressing requirement is for the development of techniques to measure the expected demands for books and journals. It is possible that the only way to treat books is to consider them as small collections in which they lose their individual identities, the size of the collections would be very dependent on the amount of detail introduced into subject classification. Even so it is likely that expected demands would have to be generated from a small amount of data about overlap in subject fields. Such information will not become readily available until library records are handled by computers.

It is also possible that, as in the example contained in this report, personal libraries should be ignored since they require too much information about the demands of individuals.

5. REFERENCES

1. C.H. Brown Scientific Serials (ACRL monograph, 16) Chicago, ACRL. 1956.

Discussion

There was a discussion on the relative merits of the two formulations of the Bradford distribution as derived by Professor Leimkuhler and Mr. Brookes:

$$R(n) = R \log (1+an-a) \qquad \text{Leimkuhler} \qquad (1)$$

$$R(n) = K \log n \qquad \text{Brookes} \qquad (2)$$

where $R(n)$ is the cumulated number of references in the first n journals, ranked in decreasing order of productivity, K, R and a are constants.

Mr. Dammers said that although Professor Leimkuhler's formula would appear to be more accurate, Mr. Brookes' version was operationally more useful in libraries.

When asked if he had investigated computer storage, Mr. Woodburn replied that he had not, and that the main difficulty of his model was that it relied on data which would never be available. He thought that the model should be regarded as a guide to visualising an hierarchical library system, with libraries higher up in the system dealing with a residual demand.

Mr. Fairthorne, comparing multi-level access and linear access, said that the former was always quicker than the latter, provided the store was large enough, because the linear access time was proportional to the size of the collection. Mr. Mackenzie remarked that speed was not always directly related to cost. Mr. Dammers said that in the computer field the access time tended to be approximately inversely proportional to the cost of storage. Mr. Buckland pointed out that there were differences between computer access and access to libraries: in linear access to computers the whole tape was searched, but in linear access to libraries there was a human element to be considered. Thus if a reader searched for an item in a departmental library and found that it was not there he would go on to search in the university library; however when he wanted the same item again he would remember that it was not in the departmental library, and he would go straight to the university library. Another factor was that he might stop searching at the departmental library level and not go on to the next stage, so that a cost would have been incurred without any resulting benefit.

Professor Leimkuhler was asked if in his model he had established a relationship between the cost of ordering an item and the cost to the user if the item was not ordered; he replied that it had not been possible to incorporate this in his model.

Mr. Fulford asked if the figures for the space saved by shelving books in different sequences, with each sequence depending on the height of the books, could be applied to any library. Professor Leimkuhler replied that his model had a general application, and he gave some figures showing the percentage increase in storage capacity gained by dividing the books into various sequences. The percentage increase gained by division into 2, 3, 4, 5 ... 10 sequences was 38%, 47% 51%, 53% ... 58% respectively. He pointed out that most libraries have an oversize sequence, which is a division into two sequences, with a corresponding increase in capacity of 38%. Mr. Mackenzie observed that it was difficult enough to educate readers to use only two sequences, and that more would merely add to their confusion.

Mr. Mackenzie said that when planning a library it was important to take into account the future developments in technology. If a building was going to stay up for

100 years great flexibility was needed so that it could be adapted to meet the changes in technology during that period. Mr. Brookes suggested that perhaps libraries should be built to last for a much shorter time, e.g. 20 years, and then pulled down for a new one to be built. Dr. Thompson said that construction techniques were so far unable to produce such a building at an economic price; Mr. Longworth added that Lancashire had tried building temporary libraries 20 years ago and they were not much cheaper than conventional ones. Dr. Urquhart said that the administration tended to become fossilized as well as the building, and this might represent a larger cost than the physical limitations imposed by the building. Mr. Fairthorne said that every administration had to adapt itself to the geography of the building which it was in.

References:

1) Leimkuhler, F.F. The Bradford distribution. Journal of Documentation, vol. 23, no. 3., 1967, p. 197-207.

2) Brookes, B.C. The derivation and application of the Bradford-Zipf distribution. Journal of Documentation, vol. 24, no. 4, 1968, p. 247-259.

SESSION 4: TECHNIQUES (Continued)

M. K. Buckland & A. Hindle	Loan policies, duplication and availability
Discussion	

LOAN POLICIES, DUPLICATION AND AVAILABILITY[x]

by

M.K. BUCKLAND and A. HINDLE
University of Lancaster

---

[x] **The work described in this paper forms part of the 'Systems analysis of a university library' project at the University of Lancaster Library and was supported by a grant from the Office of Scientific and Technical Information.**

1. INTRODUCTION

Every library has a loan policy, even if that policy is that no books may be borrowed. Libraries which do lend have widely varying loan periods from two hours to a year or more. Most libraries, especially university libraries, operate a number of loan policies simultaneously based on a variety of factors including the type of document, its value, its popularity and the status of the borrower. For example, in the University of Sussex there are five loan periods:

i. Some material is confined to the library;

ii. Some material is placed in a closed-access 'Short Loan' collection and may be borrowed for up to four hours or overnight;

iii. Some material may be borrowed for up to two days;

iv. Undergraduates may borrow the remaining material for two weeks;

v. Postgraduates and teaching staff may borrow remaining materials until the end of term.

Likewise, there is much variation in the regulations concerning renewals, reservations and recalls, in the maximum number of books that may be on loan to any one borrower at any one time and in the sanctions imposed upon borrowers who break the regulations. These variations exist between libraries and often within libraries.

Not only is there wide variation, but changes are frequently made. In view of wide variation and the central importance of loan policies to the user, the problem of choicing loan policies has been curiously neglected in the professional literature - except for rather limited discussions of the 'conflict between reference and lending'.

2. THE FACTORS INVOLVED

The main purpose of a library is to make books (and other documents) available for the clientèle it serves. Unfortunately 'availability' is a rather complex concept and considerations of economy and of the convenience of the user lead to a tangle of conflicting objectives. These can conveniently be examined by considering the various aspects of a library's loan policy.

2.1. LOAN PERIOD

The longer the loan period the more convenient it is for the borrower to use the book at leisure. However, the longer the time a book stays out on loan, the longer it is off the shelf and, thereby, less immediately available for other library users. There is a clear conflict here between the convenience of the individual borrower and the convenience of other library users who might wish to use the book. The chances that another reader will in fact wish to use the book when it is out will depend upon the level of demand for the book. It would obviously cause inconvenience if a book sought daily were to go out on loan for weeks at a time. On the other hand, if a book is rarely used (once a decade, say) then the chances of another user wanting it are small and a longer loan period can be permitted. Not only do books vary greatly in popularity, but the popularity of individual books is liable to fluctuate, although a general tendency for it to decline with time is well established.

2.2. RENEWAL OF LOAN

If a reader wishes to retain a book on loan after the expiration of the official loan period, then it is normal practice to permit one, two or even unlimited renewals unless another reader has made a reservation for that particular book. The frequency of renewal is important because it is the time that a book is absent from the shelves ('immediate availability') which matters rather than the official loan period.

2.3. RESERVATION AND RECALL

If a book is not on the shelf it can still be made available by means of a reservation and, if appropriate, recalling it from the reader who has it. To the extent to which this is an acceptable substitute for availability on the shelf, this arrangement reduces the importance of 'immediate availability' and thereby permits longer loan periods and less duplication.

Acceptability apart, this procedure of reservation is clearly unsuitable for those who are not seeking a specific title but are browsing perhaps purposefully for information on a specific topic or less purposefully for inspiration or amusement. If such a reader is browsing along the shelves, then it is clearly important that material should be on the shelves, else the reader will remain unaware of its existence and the provision of a mechanism for reservation and recall will be irrelevant. If such a reader browses in a catalogue or bibliography, then he will presumably identify particular items which he desires to inspect and his search becomes specific. In this case reservation and recall facilities become meaningful, but inspection of catalogue entries is less informative than inspection of the actual document.

2.4. DUPLICATION

So far we have been concerned with the case of there being a single copy of each different book and the effect of loan policies on the convenience of the users with respect to that book. Obviously, there is no need to enforce a short loan period to induce borrowers to return a book for the benefit of others if other copies are in fact available for them. In this manner, shortening of the loan period and duplication are alternative methods of increasing the availability of books in libraries. It must be stressed, however, that these alternatives differ in some important respects. The acquisition, processing, and storage of each additional copy of a book costs money and labour which could well be put to other uses, such as another different book or any other library service, and to this extent duplication is undesirable. The shortening of the loan period and the restricting of renewals, however, are likely to cause additional inconvenience to the borrower and to this extent are undesirable. Furthermore, the shortening of loan periods is likely to involve additional administrative expenditure.

2. 5. ADMINISTRATION

Apart from considerations of availability and the cost of duplication, the cost of administering a loan policy must also be considered. A short loan period might be expected to result in a larger number of renewals. A low level of immediate availability (whether by long loan periods or limitation of duplication) is likely to be associated with larger number of reservations and recalls. A policy decision to maintain a borrowers file either as a service to readers or as a means of enforcing a limitation on borrowing will also involve additional expenditure.

A recent trend in British university libraries towards an "until the end of term" type of loan period seems to have been at least partly the result of a desire to economise on service desk staff by avoiding the necessity for a file arranged by date and reducing the number of renewals.

In addition, distinctions between borrowers and especially between books (e. g. when the more popular books are treated differently from less popular books) are likely to involve extra administrative expense.

However, although the loan policies will affect the number of transactions of various kinds, the unit costs will depend very much on the details of the issue system employed, which is outside the scope of this paper.

2. 6. STATUS OF BORROWER

In some libraries the loan policies give privileged rights to some users - not just in exceptional circumstances, but as a normal practice. In university libraries, for example, teaching staff are almost invariably subject to a longer loan period and a higher limitation on the number of books allowed out at any given time. One explanation of this, is that books used by students are in heavy demand and that, therefore, a shorter period of loan is more appropriate. Another explanation is that teaching staff are more important.

2. 7. FINES

The principal method of enforcing loan policies is the charging of fines which also vary considerably from library to library. The assumption is that fines are necessary to ensure compliance with regulations. Unfortunately, the relationship between fines and borrower behaviour has not, to our knowledge, been subject to serious investigation and hence this assumption must remain doubtful.

2. 8. SUMMARY

It will be clear from the preceding sections that wide variations in loan and duplication policies reflect complicated relationships involving a number of conflicting objectives. Any rational loan policy must be a considered compromise.

During the University of Lancaster Library's recent research into the application of operational research techniques to library problems, the clarification and measurement of these relationships have been examined and this spring the Librarian requested the authors to appraise this library's loan policies and to prepare recommendations for the university's Library Committee concerning what changes, if any, should be made. The rest of this paper outlines this work.

3. THE BASIC RELATIONSHIP

3.1. A SINGLE TITLE

The most convenient measure of the availability of a library book is the proportion of times that it is immediately available on the shelf when sought. This can be expressed as a percentage, thus:

$$\% \text{ immediate availability} = \frac{\text{no. of times immediately available when sought} \times 100}{\text{no. of times shought}}$$

Apart from the number of copies held, the two critical factors determining the immediate availability are:

(i) the frequency with which the book is sought (its 'popularity'); and

(ii) the length of time it is off the shelves when in use.

For any given level of demand, the book will become more often available if users can be induced to return it to the shelves more quickly. This is the one justification for 'reference' libraries where no borrowing is permitted and so the time a book is absent from the shelves is a matter of hours not weeks.

For any given pattern of return times (the time the book is off the shelf) the availability will depend upon the level of demand. A rarely requested book is more likely to be available when sought than a book which is frequently requested.

Obviously the number of copies held is also very important in determining whether a copy will be available when sought.

The chances of an unsuccessful library user making a reservation is also important because a reservation will delay the book's return to the shelves.

3.2. MONTE CARLO SIMULATION

In exploring these relationships we have used a technique known as Monte Carlo simulation. Other workers have employed queuing theory and other stochastic process models. For an extended discussion of the application of these formulae to library problems (especially loan policies and duplication) reference should be made to the recent book[1] by Professor P. M. Morse entitled Library effectiveness: a systems approach.

The simulation approach has certain advantages over the more strictly mathematical treatments. For example, it allows greater freedom in the

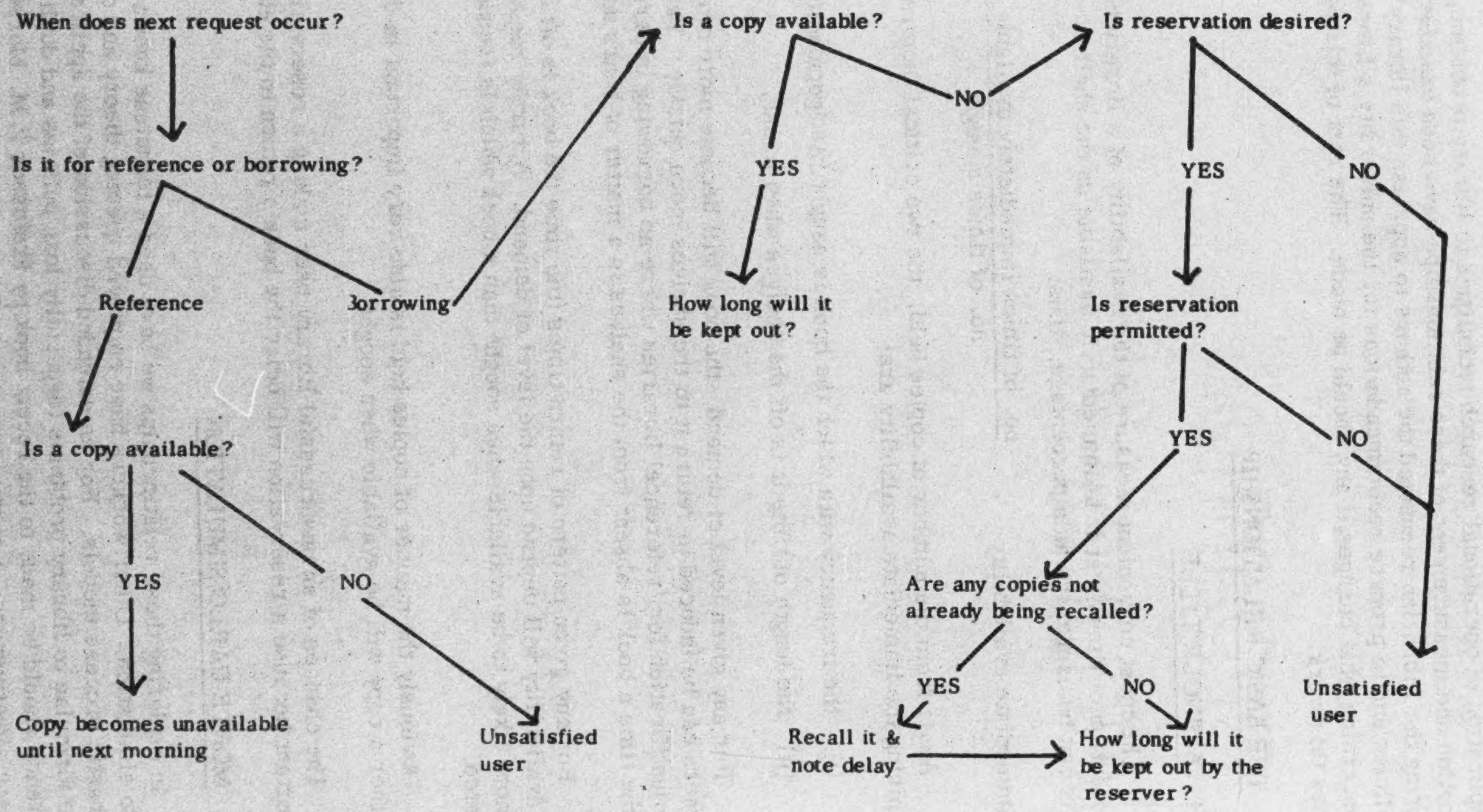

When does next request occur?
Is it for reference or borrowing?
Reference
Borrowing
Is a copy available?
YES
NO
Copy becomes unavailable until next morning
Unsatisfied user
Is a copy available?
YES
NO
How long will it be kept out?
Is reservation desired?
YES
NO
Is reservation permitted?
YES
NO
Are any copies not already being recalled?
YES
NO
Recall it & note delay
How long will it be kept out by the reserver?
Unsatisfied user

FIGURE 1: Description of the borrowing process

assumptions that can be made and hence has greater facility in handling complex relationships. A description of the borrowing process is fed into the computer. The description currently in use is shown in Figure 1 on page 6. In addition, some numerical data is also required, which defines, for example, the number of copies, ratio of borrowing to reference (in-library) use and the probability that an unsatisfied user will make a reservation. The computer performs the borrowing process several thousands of times, keeps a record of significant events, and then reports the result.

In real life, accurate predictions are often not possible. For example, it is not possible to state with confidence whether the next reader will use a book in the library or borrow it. Also it is not possible to state precisely in advance how soon the book will be returned to the shelves. However, data can be collected which shows what the overall pattern is in the long run. Similarly, the computer determines such problems by picking numbers at random - as if by rolling dice (hence the name Monte Carlo simulation). Each possible number will have had a particular answer associated with it and this association of numbers and answers will have been pre-arranged so that in the long term the overall pattern will correspond to the overall pattern of the system being simulated.

In this manner it is possible to simulate situations of great complexity so long as the situation can be described in logical and probabilistic terms. In order to assess the effect of a change in a situation, a simulation of the actual current situation is designed and carried out. The simulation results will, if the simulation is valid, correspond to the actual current results. Another simulation is then performed incorporating the proposed change and the results indicate in detail the likely consequences of the proposed change. An important feature of this technique is that it is possible to examine how sensitive the system being simulated is to particular changes. For a fuller description of Monte Carlo simulation the interested reader should refer to the appropriate textbooks, such as K. D. Tocher: The art of simulation.[2] Before describing in detail the results of our simulations, two other relationships need to be examined.

3. 3. THE RELATIONSHIP BETWEEN ACTUAL LENGTH OF LOAN AND OFFICIAL LOAN PERIOD

Regulations concerning loan periods are, in effect, a control device at the disposal of the librarian for influencing the movement and thereby the availability of the books in his charge. It would, therefore, seem important to examine the nature of this influence.

At first sight the factors which seem likely to affect the length of time are both numerous and varied: the subject, the level and the type of the book; the subject background, the work habits and the motivation of the borrower; the thousand and one possible distractions which might affect his behaviour (including recall notices); and the official loan period. In a situation where the factors involved are of such complexity, one would expect books to be returned as if at random - of all books borrowed on a given day a small percentage of those still out being returned each day - with perhaps a small peak at the expiration of the loan period. In order to test this hypothesis data relating to a variety of loan policies in a number of university libraries were examined as follows:

One week	Manchester	
Two weeks	Michigan,[3] Strathclyde, Sussex	
Four weeks	Manchester, Michigan,[3] Strathclyde	
Ten weeks	Strathclyde	
End of term	Lancaster)	An entire year's discharged issue slips permitted a variety of analyses to be made.
End of year	Lancaster)	

At none of these libraries is a limitation on the number of books allowed out rigorously enforced. (We have no information about Michigan in this respect).

The pattern which emerged consistently is that there is a very strong tendency for books to be returned or renewed at the expiration of the official loan period - whatever the length of the loan period may be. Only with comparatively long loan periods - longer than four weeks - is a substantial amount of material returned before the due date and here the expected negative exponential pattern emerges. (See Fig. 2 on page 9).

3.5. LOAN PERIODS AND RENEWALS

A further analysis was made to see whether shorter loan periods were associated with increased renewals. The results (Table 1 on page 10) show that the proportion of books renewed one or more times varies little over a range of loan periods.

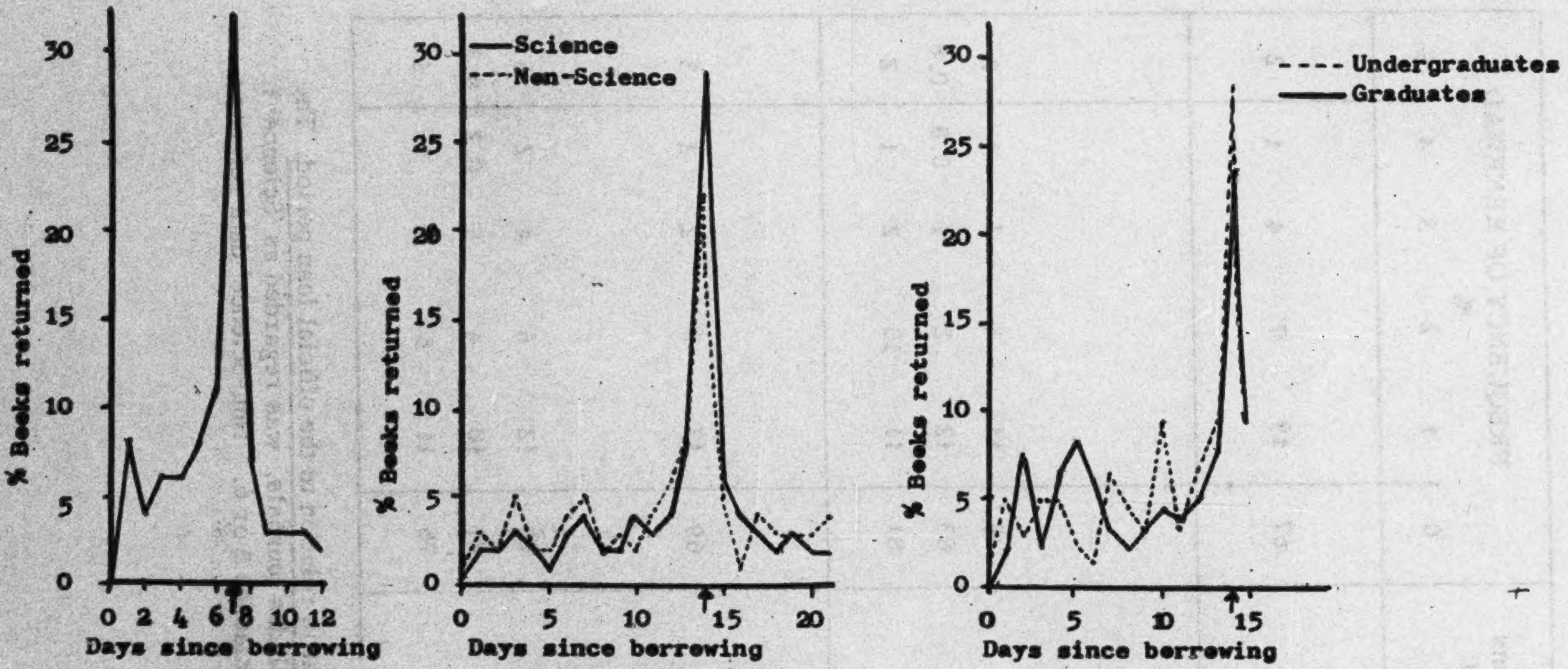
With two exceptions the results appear to be comparable - the reader presents the book at the Service Desk and requests a renewal. The two exceptions are the teaching staff at Manchester where the loans are automatically renewed on dates determined by the borrower's surname and undergraduate two-week borrowing at Sussex where only one renewal is permitted and that requires rather more effort on the user's part.

4. APPLICATION

4.1. SIMULATION

Our present simulation model, described above, requires the following information:

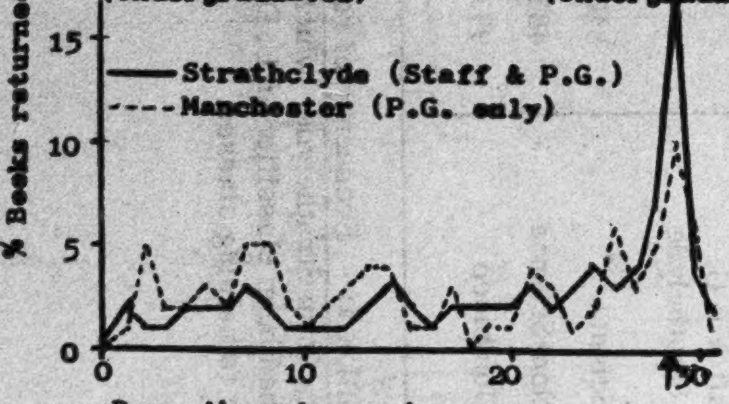
1. The number of demands to be simulated;
2. The number of copies of the book;
3. The maximum allowable number of reservations per book;
4. The pattern of demand as expressed by the intervals between requests;
5. The ratio of borrowing to reference use;



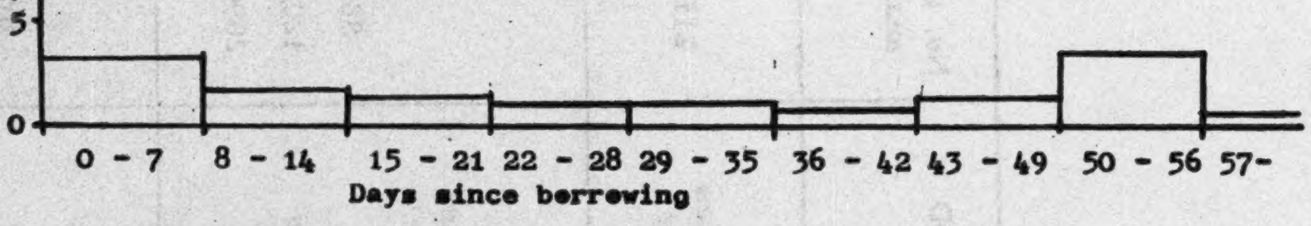
1 WEEK - MANCHESTER
(Undergraduates)

2 WEEKS - STRATHCLYDE
(Undergraduates)

2 WEEKS - MICHIGAN



4 WEEKS - MANCHESTER & STRATHCLYDE



8 WEEKS: LANCASTER (Undergraduates): Material borrowed 50-56 days before the date at the end of term when all loans were due back.

FIGURE 2: HOW LONG BOOKS ARE KEPT OUT. The number of books (%) returned on successive days since they were borrowed. Data from Manchester and Strathclyde exclude renewals and items reserved by other readers.

LOAN PERIOD	No. of loans analysed	FREQUENCY OF RENEWAL %					
		0	1	2	3	4	5+
1 WEEK:							
Undergraduates at Manchester	2115	67	19	7	4	1	2
2 WEEKS:							
Undergraduates at Strathclyde							
- Science	888	78	11	3	3	1	2
- Non-Science	1208	83	12	2	1	0.8	0.8
- Combined	2096	81	11	13	2	1	2
4 WEEKS:							
P.G. at Manchester	307	69	17	6	2	2	3
Staff & P.G. at Strathclyde							
- Science	510	70	12	6	4	2	6
- Non-Science	483	81	10	4	5	0.2	0.2
- Combined	993	76	11	5	4	1	3

TABLE 1: Frequency of renewal in relation to the official loan period. The data from Strathclyde, which excludes journals, was regarded as 'Science' if it had been classified in U.D.C. classes 5 or 6. 'Non-Science' data refers to the remaining classes.

6. A loan policy as defined by the return times;
7. The proportion of unsatisfied users who make reservations;
8. The delays involved in recalls.

The computer reports the following results:

For reference:

a. The number of reference demands made;
b. The percentage of satisfied reference demands.

For loan:

a. The number of loan demands made;
b. The percentage of satisfied loan demands;
c. The level of immediate availability (i. e. the percentage of demands satisfied immediately);
d. The pattern of delays experienced following reservation.

The loanable stock of the University of Lancaster Library (excluding the closed access Short Loan collection of very heavily used text books) was then analysed by examining borrowing histories. The result is shown in Table 2. This pattern lends support to the hypothesis that the library's books are subject to demands approximating the Zipf's law distribution.[4]

No. of issues per annum	0	1	2	3	4	5	6	7	8	9	10
Frequency	375	168	103	43	40	15	6	2	1	1	1
%	50	22	14	6	5	2	1	-	-	-	-

TABLE 2: Analysis of borrowing histories, 1967-1968. Out of a random sample of 876, 757 were analysed and 119 were either missing or on loan.

For simulation experiments six popularity classes were considered: popularity being measured by the number of demands per time period (T) as follows:

Popularity class	A	B	C	D	E	F
No. of borrowing demands per time period T	4½ +	3½ - 4½	2½ - 3½	1½ - 2½	½ - 1½	0 - ½

By attempting to fit varous demand distributions to the borrowing history data the proportion of books in each popularity class was determined, where T is one year for the Lancaster collection.

The simulation results gave the immediate availability to be associated with each combination of loan policy and popularity level. Subsequent calculations determinted various parameters for the collection as a whole (as described below).

Table 3 on page 12 illustrates the results of a particular simulation experiment.

POPULARITY CLASS	ONE COPY				TWO COPIES				THREE COPIES			
	Loan Policy				Loan Policy				Loan Policy			
	i	ii	iii	iv	i	ii	iii	iv	i	ii	iii	iv
A	91	79.	52	37	100	98	84	66	100	100	97	86
B	94	86	62	44	100	99	91	77	100	100	99	93
C	98	94	72	56	100	100	97	87	100	100	100	98
D	99	98	82	68	100	100	99	84	100	100	100	100
E	100	100	97	85	100	100	100	100	100	100	100	100

TABLE 3: Simulation results showing immediate availability (%) for 60 trials: 5 popularity classes A-E x 4 loan policies, i-iv x 3 levels of duplication. The four loan policies correspond approximately to loan periods of one week, two weeks, five weeks and ten weeks respectively.

4.2. ANALYSIS

In examining this matter more closely three measures of effectiveness were employed.

"Immediate Availability": The probability that a request for a given book will be satisfied immediately, i.e. the probability that a book is on the shelves.

"Satisfaction Level": In a given time period the proportion of demands immediately satisfied. All demand, that is, not just demands for one individual title only. This appears to be the most useful single measure of library effectiveness. It is not the same as mean immediate availability.

"Collection Bias": Commonly the most strongly recommended books are removed to a closed access reserve collection and many of the other more popular books will be out on loan. Consequently a reader who seeks for material on a given subject is faced by an array which systematically tends towards the least popular, the least recommended and the most shunned material. We define this tendency as a "negative bias" and one convenient measure of it is the proportion of the 10% most popular books which are absent from the shelves.

We consider that, subject to considerations of user convenience, a good library should have a high Satisfaction Level and a low Collection Bias. Beyond this, however, it is very difficult to justify any specific Satisfaction Level or Collection Bias until more is known about the effect of each on the library behaviour of users. Clearly, a reference-only library would rate well on both measures, but this entirely neglects out-of-binding use, which is generally regarded as desirable.

Analysis of the simulation results in relation to the University of Lancaster Library resulted in an estimated Collection Bias of 45%. These estimates were derived as follows. Given that a demand distribution for the library has been estimated, it is possible to calculate the probability that a demand (occurring at random) will be for a book within a particular popularity class. If this probability is multiplied by the immediate availability of the popularity class (obtained from the simulation results), then the product is the probability that the next demand will be for a book within a particular popularity and satisfied. The Satisfaction Level for the collection as a whole is the sum of these products for the six popularity classes. The Collection Bias can be calculated directly from the immediate availability of the most popular 10 per cent of the collection. It is simply 1 - immediate availability. A check on these estimates was made by comparing a prediction of the number of books out on loan at the mid-point of the 1967-68 session (to which all data refer) derived from the above analysis with a separate analysis based on issue records and discharged issue slips.

The loan simulation was then employed to predict the likely consequences of various different policies. Three main courses of action emerged. One was that there should be no change in loan policy, but selective duplication. Another was that there should be no change in loan policies for staff or postgraduates, but a two week loan period for undergraduates. However, since the problem stems from a small proportion of very popular books, the policy finally adopted was to restrict the loan period for these books only, regardless of the status of the borrower. One of the advantages of a variable loan policy of this type is that a library can be adjusted to any reasonable Satisfaction level or Collection bias by varying the definitions of popularity. Furthermore, with any given definitions, the library would be self-adjusting as the pattern of demand changes. The monitoring of the library stock in order to assess the popularity provides the basic data necessary for the management of the loan policy and also yields information of value for efficient selective duplication.

The policy approved by the University is as follows:

(a) For the 10% most popular books the loan period should be one week (- an estimated 70% of issues);

(b) For all other books the loan period should be until the end of term.

Using the methodology outlined above, we estimate that this policy will give a Satisfaction level of 85% (an increase by 40%) and a Collection bias of 10%.

4.3. IDENTIFICATION OF POPULAR BOOKS

The only data which is readily available concerning the use of individual documents are the dates on the date label. It has been shown by Fussler and Simon[5] and by Lister[6] that records of past use are reliable predictors of future use. Certainly the record of borrowing is an incomplete record of total use, but it is the critically important part of total use so far as availability is concerned. The problem has four parts, which we proposed to tackle as follows:

(i) Existing stock: The date labels of the entire loanable stock will be examined this summer: borrowed books as they are returned, the rest by going along the shelves during the vacation. Books subject to a seven day loan will be marked on the spine and the date label.

(ii) Falling popularity: At intervals, the shelves will be inspected and the date labels of volumes market "seven-day loan" will be inspected. If use has declined the markings will be removed and the book reverts to the longer loan period. Some "seven-day loan" books will be out on loan and they will be assumed to have stayed popular.

(iii) Rising popularity: This is likely to be only a small problem because of the well established tendency for the popularity of books to decline with time. Consequently repeated checking of the entire stock is likely to prove unjustified. Instead the staff of the Service Desk will be authorised to make any book subject to "seven-day loan". The fact that a reservation has been made for a book is in itself excellent prima facie evidence that that book is in heavy demand. In addition, occasional spot checks will be made on "non-popular" books being returned from loan.

(iv) New books: Those who recommend titles for purchase will be asked to mark an appropriate box on the Suggestion card if they expect a book to be popular. The subject specialists who classify books will also indicate books which are likely to be heavily used. An incorrect prediction either way will be corrected by the continuous monitoring of (ii) and (iii) above.

This is, of course, only one possible way. All this information, except concerning new books, could readily be produced as the by-product of an automated issue system. An alternative approach for either manual or automated systems would be to determine the period of loan at the time of issue - without having the books processed into pre-determined loan categories.

4.4. IMPLEMENTATION

As stated above any rational loan must be a considered compromise between a variety of conflicting objectives. The new policy at Lancaster is intended to achieve the following:

(i) 10% of stock to be subject to one week loan;
(ii) 70% of issues to be subject to one week loan;
(iii) 85% Satisfaction Level;
(iv) 10% of most popular books absent from shelves;

And it was based on:

(v) the distribution of demand;
(vi) the level of demand;
(vii) the size of stock;
(viii) the level of duplication, and
(ix) the size and rôle of the Short Loan collection

which existed in 1967-1968.

The problem of implementation is essentially one of devising simple, inexpensive procedures which will maintain the standards of service implied by the choice of loan policies.

REFERENCES

1. MORSE, P. M., Library effectiveness: a systems approach. Boston, Mass., MIT Press. 1968.

2. TOCHER, K. D., The art of simulation. London, English Universities Press. 1964.

3. Published data from BURKHALTER, B. R. and RACE, P. A. An analysis of renewals, overdues and other factors influencing the optimal charge-out period. In BURKHALTER, B., ed. Case studies in library systems analysis. Metuchen, N. J., Scarecrow Press. 1968. pp. 11-33.

4. BUCKLAND, M. K. and HINDLE, A. Library Zipf. Journal of Documentation. 25(1) March, 1969. pp. 52-57.

5. FUSSLER, H. H. and SIMON, J. L. Patterns in the use of books in large research libraries. Chicago, University of Chicago Library. 1961.

6. LISTER, W. C. Least cost decision rules for the selection of library materials for compact storage. Ph. D. thesis, Purdue University. 1967. (Clearing house report PB 174 441).

Discussion

Mr. Duchesne asked about the cost entailed in shortening the loan period, and also if borrowers would borrow more books with a shorter loan period. Mr. Mackenzie replied that the available evidence indicated that the number of books borrowed per reader would increase. The extra costs involved were in redesigning the mechanics of the loan system, in staff time for sending out the greater number of overdues which would be required, and in extra staff time of 6 man-weeks per year needed to monitor the stock and to make sure that books were in the correct category; there was also an extra cost of 0.5d. per issue for a more sophisticated type of stationery which was needed. Mr. Buckland said that one factor which would make the total demand go up in the future would be more students in residence; also if availability were increased this might lead to a greater demand on the library because readers would realize that there was more chance of finding what they wanted on the shelves, and a self-reinforcing situation would be created.

Dr. Urquhart wondered if it would have been better to spend the money on duplicates to increase availability, rather than on redesigning the system. Dr. Hindle replied that the system for ordering duplicates could be improved; there seemed to be the usual lack of communication between the library and the departments about books which were popular; he also pointed out that a variable loan policy did not preclude selective duplication. Mr. Mackenzie added that about 40% of the recommendations for books to be placed in the closed access short-loan collection were dubious because the books were hardly used; but that the library spent £2,000 - £3,000 per year on duplication which was initiated by the library from evidence gained at the service desk; consequently he doubted the ability of the departmental staff to forecast demand. When asked at what time the decision on the category of new books would be taken, Mr. Buckland replied that this would be done on ordering or by subject specialists when they classified the books. Mr. Mackenzie observed that no serious penalty would be incurred by the reader if a book were wrongly placed in the popular category, since six months later lightly used books would be down-graded. Mr. Snape said that for particular titles some public libraries had employed a saturation technique and duplicated until there were always copies of those titles on the shelves. Mr. Mackenzie said that this had been proposed on a small scale as part of the research project, but it would be very expensive to implement as a standard practice.

Dr. Thompson asked if Lancaster's acquisition policy took into account the fact that students bought some of their own books and whether the library deliberately avoided bying strongly recommended texts. Mr. Mackenzie replied that departments had different policies on the purchase of books by students: some departments issued lists to their students stating which books they had to buy and which they should get in the library; the library would probably buy one copy of the former and many copies of the latter.

Mr. Buckland said that the present satisfaction level was 60% and they had initially aimed to increase this to 80%, but this had only been an intuitive guess; he wondered what the optimum level should be, and what effect the satisfaction level had on user behaviour. Dr. Urquhart said that at the Science Museum library 20% of the stock used to be on loan when it was wanted by users, and at the NLL they aimed to reduce this figure to 10%; but they had no positive measure of what the figure ought to be.

Mr. Buckland said that many libraries relied on limiting the number of books which readers were allowed to have out at any one time as a means of increasing the satisfaction level. This seemed rather a doubtful practice because it only affected the "good" user, i.e. the person who used the library a lot, and it was not known if such users tended to return the popular or the unpopular books. Mr. Vickery asked if there were any data on the relationship between the number of books a reader had out, and the number they were allowed to have. Mr. Buckland replied that he had not, because, of the libraries they had examined, all put a limit on the number of books readers were allowed to have out, but these limits were not enforced because the additional staffing cost was too great. Mr. Fulford asked if the cost of raising the satisfaction level increased as the starting point got higher. Mr. Buckland replied that it did, but he had no precise figures, since the indirect cost to the institution as a whole was impossible to calculate.

Mr. Buckland said that a number of studies had suggested that there was a correlation between the use made of a library and academic achievement in examinations. Therefore if students who were good academically choose to allocate their time to using the library, then it could be hypothesised that those who did not use the library very much should be encouraged to use it more: if this could be done by increasing the satisfaction level, especially for new students, then the library would be performing a useful service. Dr. Urquhart asked if there were figures available which gave the percentage of students who never borrowed a book. Mr. Buckland replied that there were no figures for Lancaster but he thought that the percentage must be high; the Nuffield Survey at Leeds found that 26% of undergraduates never borrowed a book. Dr. Urquhart said that Leeds was not really comparable with Lancaster.

Dr. Urquhart said that one way of increasing the satisfaction level in a particular section of the collection would have been not to bind the periodicals into large volumes. Mr. Mackenzie said that the study was basically concerned with undergraduate usage and therefore periodicals were not considered. Mr. Buckland added that it was cheaper to bind periodicals into volumes than to bind each part separately. Mr. Mackenzie pointed out that it was easier to steal a separate issue of Physical Review than it was to steal a bound volume of it.

Dr. Thompson asked if there was any information available on the differences between the borrowing patterns of staff and students. Mr. Mackenzie regretted that there was very little information on this. Dr. Urquhart suggested an alternative loan period of one week for all loans, coupled with encouragement to readers to fill in reservation forms and an understanding that books would not be recalled until the end of term unless someone else wanted them. Mr. Mackenzie said that this was in essence the existing system, but only about 10% of readers filled in reservation forms when they wanted books which were not on the shelves. Mr. Brookes wondered what effect the new loan policy would have on student buying of books. Mr. Mackenzie replied that the new policy had not been implemented yet, but it would be interesting to see if it had any effect. Mr. Buckland said that some studies indicated that readers who borrowed a lot from public libraries also tended to buy books.

SESSION 5: DATA COLLECTION AND EVALUATION

R. M. Duchesne	Library management information from computer-aided library systems
M.G. Ford	Data collection and feedback
A. Hindle	Models and measures for non-profit making services
Discussion	

LIBRARY MANAGEMENT INFORMATION FROM COMPUTER-AIDED LIBRARY SYSTEMS

Contents

	Page
1. Abstract	2
2. Management Information (MI)	
2.1 The Librarian as Manager	3
2.2 Information Needs of the Librarian/Manager	4
3. Computer-aided Library Systems	
3.1 MI: the Role of the Computer	6
3.2 MI in presently reported systems	8
3.3 Towards conditions favouring design of systems giving more effective MI	9

Appendices:

A. Some Desirable MI for the Librarian/Manager

B. Some Suggestions for Byproduct to be produced by Computer-aided Library Systems

R.M. Duchesne
University of Birmingham
June 1969

1. ABSTRACT

Librarians are managers; as managers they need management information. Appendix A provides a sample list of desirable MI.

The introduction of computer-aided library systems (e.g. ordering, cataloguing systems) enables libraries - if the systems are designed with this in mind - to collect and process more MI than can feasibly be collected in existing manual systems. Appendix B suggests MI which may be produced by present-day computer-aided systems. The possibilities for collecting MI of this type increase as more processes are automated, especially if machine processes are integrated, i.e. dovetail and make use of common files. When it becomes feasible for users to communicate directly (say, via consoles) with the machine system, even greater possibilities will arise.

While computer-aided systems offer MI possibilities, present reporting suggests that neither librarians nor computer men really know what data to collect, analyse and display for MI purposes. The relatively simple Appendix B type of information will be of considerable use; the next step on from this type of information is to make use of an operational research (OR) approach. To take this step forward the applicability of the OR approach in the library context needs to be more fully demonstrated; also the library profession, including its systems analysts, need to become better acquainted with the OR approach. Research projects employing OR techniques in libraries would help demonstrate the applicability of OR in libraries. Acquaintance of the library profession with OR would be improved by short courses on the subject; by the production of a monograph on library systems analysis, including treatment of the application of OR techniques to library problems; by more adequate treatment of OR in formal full-time professional education.

2. MANAGEMENT INFORMATION (MI)

2.1 The Librarian as Manager

Brech[1], who is a widely quoted authority, has defined managment in the following way:

> "A social process entailing responsibility for the effective planning and regulation of the operations of an enterprise, in fulfilment of a given purpose or task, such responsibility involving:
>
> (a) The installation and maintenance of proper procedures to ensure adherence to plans.
>
> (b) The guidance, integration and supervision of the personnel composing the enterprise and carrying out its operation".

He goes on to define direction (concerned with objectives and policy), organisation, and administration (concerned with procedures) as parts of management.

Such a definition is descriptive of the main tasks undertaken by individuals in charge of libraries, particularly libraries of any size. It is therefore odd that librarians, especially in UK, have not shown enthusiasm in identifying themselves as managers. It is surprising, for instance, how few sizeable monographs there are, devoted to library management and library administration. Even some of these books avoid stating outright that the librarian having charge of a library is a manager and must endeavour to be effective as such. For instance, the following statement is a diffident version of the latter thesis, especially in a book on library management: "although the foundationstone of librarianship is bibliography, the profession also includes many elements common to business and industry. Business executives and industrial engineers have long known the value of scientific management".[2]

---

1. BRECH, E.F.L. (ed.). The principles and practice of management. Longmans, 1953, (p.17).

2. DOUGHERTY, R.M. and HEINRITZ, F.J. Scientific management of library operations. Scarecrow, 1966. (Statement in preface)

One can see many reasons why librarians have been diffident in identifying themselves as managers. In British academic institutions, there is a tendency to regard administration and management as activities for the second-rate mind and therefore as of less importance and status than teaching and research. A certain defensiveness within the profession has probably also encouraged the view that librarians know about libraries and do not have anything of much importance to learn from managers of other institutions.

In fact the librarian in charge of a library is just as much a manager as - assuming similar sizes of organisation - the factory manager in an industrial enterprise, the manager of one of a chain of department stores, a bank manager. The form of authority to which each type of manager is responsible differs, but each has the responsibility for staff, buildings, equipment, financial expenditure and has the task of managing those resources to produce some form of product or service. In the case of libraries, the resources can be on a very substantial scale. Such is the case for the major national libraries like the Library of Congress or the British Museum. There are also very large university libraries; Harvard, for instance, has an annual budget of over $8 million and a staff of more than 600. The average U.K. university library and the average U.K. public library authority are of course not nearly so large as this, but with average annual expenditures of some £160,000 (1963/64)[1] and £110,000 (1967/68)[2] respectively - even the average librarian in these types of library commands an impressive scale of resources.

2.2 Information Needs of the Librarian/Manager

As a manager, the librarian needs information on which to base decisions. Indeed management has been described[3] as the "process of converting information into action. The conversion process we call decision making". The number and complexity of decisions required of the librarian is large, since:

---

1. UNIVERSITY GRANTS COMMITTEE. Report of the committee on libraries. HMSO, 1967, page 155.

2. INSTITUTE OF MUNICIPAL TREASURERS AND ACCOUNTANTS. Public library statistics 1967-68. IMTA and Society of County Treasurers, 1969, page 6.

3. GREENBERGER, M. (ed.). Computers and the world of the future. MIT, 1962, page 37.

(1) The library needs to adapt to a changing environment: for instance, changing user needs, changing patterns and volume of publication. The needs of research workers, particularly, tend to be subject to rapid change.

(2) Many decisions are complex: for instance, how to allocate the budget - how much to spend on staff, how much on material; how much on books as against periodicals.

(3) Wrong decisions may be costly, and may be difficult or impossible to rectify. For instance, a library building embodying one concept of the library cannot easily be changed if not found satisfactory in practice.

(4) As libraries grow larger, as the pace of technological change increases, as the rate of publication increases, the librarian's management task is becoming more onerous - and his need for management information is growing.

To cite one instance: with the growing volume of publication libraries face a greater selection task - there is more to scan and select. Add to this the fact that most individual libraries' budgets are not growing as fast as world publication, and the task of selection is becoming not only greater but more necessary.

What sort of MI does the librarian need? Appendix A illustrates some essential information. For a more detailed review of information required for management purposes, reference may be made to McDIARMID.[1]

---

1. McDIARMID, E.W. The library survey; problems and methods. ALA, 1940. The reader should not be put off either by the date or the title.

3. COMPUTER-AIDED LIBRARY SYSTEMS

3.1 MI: the role of the computer

From the sample list of MI in Appendix A it is clear that much MI will remain outside computer systems for a long time, possibly for ever - for instance, information concerning organisation structure, job specifications, staff and procedure manual information.

With regard to MI, then, the role of the computer is a limited one. Firstly it will not, forseeably, supply all the librarian/manager's formal information needs. Secondly - it goes without saying - it will never supply the sort of informal MI which can only be obtained by personal contacts. The library, after all, exists to serve human beings and is staffed by human beings; it must be managed by a human being - the limited role of the computer arises partly from these simple facts. There is no chance that the computer will dehumanise library management; neither is there any chance that it (or the computer man) will replace the librarian/manager, take decisions for him, or reduce his sphere of influence. The librarian/manager himself will always have the final say concerning all matters under his charge, including the design[1] of his MI system, what MI reports he is to receive, and what action to take on these reports.

Having stressed the limited role of the computer and that the librarian/manager will never obtain all his MI from the computer, even if all the procedures in his library are computer-aided - the question still remains - what MI can he expect the computer to supply?

At present relatively few libraries have introduced automation on any scale. For those without computer-aided systems, the main MI role of the computer is to process survey-type information. For those with computer-aided systems (e.g. computer-aided cataloguing, circulation systems), the possibilities are considerably greater, since the computer may be used to produce MI information as a byproduct. Appendix B gives some indication of the type of byproduct information which may be collected. Some of this may warrant routine reports, some might be presented in reports produced on request.[2]

---

1. LOSTY, P.A. Designing a management information system (MIS). Computer Bulletin 13 (5) May 1969, 142-147. (Page 143).

2. For discussion of types of report which may be produced and who may call for them, see:
CLOOT, P.L. Management information systems - can computers help? Computer Bulletin 11 (4) March 1968, 276-281. (Page 277).

The systems postulated in Appendix B are relatively simple self-contained batch-processing systems: batch-processing selection data, order, cataloguing, circulation systems. The systems are those which general libraries introducing computer-aided systems are most likely to develop first, and those from which MI of general validity can be readily extracted. SDI and IR systems have been excluded on the grounds that to extract MI of general validity they must be used by a large proportion of the library's users. SDI and IR systems on such a scale are not so likely to be economically feasible for general libraries as those listed.

Relatively self-contained systems are chosen because it is to be expected that most libraries introducing computer-aided systems will start by creating some relatively self-contained system of the general type indicated. As automation proceeds and integrated systems are developed, e.g.

selection/order
order/cataloguing
cataloguing/circulation

the possibilities for byproduct MI increase. For instance, an integrated selection/order system can keep statistics of the delay between selection and ordering.

Possibilities for collecting MI will be further increased by:

1. Development of 'total' systems (e.g. covering selection/order/cataloguing/circulation).
2. Direct user communication with the machine, (e.g. via consoles) - when this becomes feasible. This is looking some way into the future; a computer system with which the user could communicate in conversational mode offers prospects of building individuals' profiles, noting relatively precisely what information they use and their mode of finding relevant information. Licklider[1] has described such a system.

---

[1] LICKLIDER, J.C.R. Libraries of the future. MIT, 1965.

3.2 MI in presently reported systems

There are at present relatively few operational computer-aided library systems. Taking a cross-section[1] of reported systems, the projected total systems show awareness of potentialities. To quote from the Council on Library Resources report (page 23):

"Possibly one of the most important consequences of automation will be to provide capability for maintaining use history and for implementing measurements on a sampling basis in order to install a good system of quality control in the library's operations".

Similarly, in the Intrex report, Morse (in Appendix N of the report) stresses the need for MI, puts forward some "typical probabilistic models of library operations, and states that "... the introduction of data-processing equipment in library operations will make it easier to amass the data [needed for use of these models]" - page 228.

---

[1] Projected total systems

COUNCIL ON LIBRARY RESOURCES. Automation and the Library of Congress, Library of Congress, 1963.

OVERHAGE, C.F.J. (ed.) Intrex; report of a planning conference. MIT, 1965.

Order system

GROSE, M.W. The Newcastle University Library order system
in COX, N.S.M. and GROSE, M.W.
Organisation and handling of bibliographic records by computer.
Oriel, 1967, p.158-159.

Cataloguing

BREGZIS, R. The Ontario new university library project. College and Research Libraries, v.26. November 1965, 495-508.

Circulation

WOODS, R.G. Use of an ICT 1907 computer in Southampton University: Report No. 3.
Program v.2 (1) April 1968, p.30-33.

Note: More reports than these have been examined - e.g. that on University of Hawaii circulation system - but the above are sufficient for the purpose in hand.

Morse and (in Appendix R to the report) Raymond in the Intrex report are - in the cross-section of reports examined - by far the most articulate concerning the need for MI and how to use it when obtained.

Grose notes eight tabulations to be produced by a program under development at the time of his report. Bregzis merely states that a wide range of information can be produced as necessary, while Woods makes no mention of MI though the Southampton system does in fact accumulate various counts, e.g. of loans by type of borrower.

The overall impression and conclusions to be drawn are well put by Morse[1];

"... neither the computer experts nor the librarian (for different reasons) really know what data would be useful for the librarian to have collected, analyzed, and displayed, so he can make decisions with some knowledge of what the decision implies. What is needed before the computer designs are frozen is for models, of the sort developed in this book, to be played with, to see which of them could be useful and to see what data are needed and in what form, in order that both models and computers can be used most effectively by the librarian".

3.3 Towards conditions favouring design of systems giving more effective MI.

The operations research (OR) approach followed by Morse and others, notably Leimkuhler[2], seems to offer great possibilities and to point the way forward for library MI - with certain reservations:

(1) The Need to test OR theories in practice.
Most of the literature concerning OR in libraries contains statements such as these:

---

[1] MORSE, P.M. Library effectiveness: a systems approach. MIT, 1968 p.141-142.

[2] See, for example:
LEIMKUHLER, F.F. Systems analysis in university libraries. College and Research Libraries, v.27 (1) January 1966, 13-18.

> "It is unlikely that they [the author's mathematical models] are precise indicators of actual performance because of the assumptions required in their development, as, for example that the demand rate for a book is independent of its availability or the loan policy".[1]

> "The book is experimental in another sense; it is carrying theory well beyond the point where it can be solidly backed up by presently available data ... most of the models discussed here ... are backed by the flimsiest of corroborative measurements".[2]

There is need to test OR theoretical models by application in real-life library situations. Probably the quickest way to bring about such application is through research projects. The object of such projects should be to apply OR techniques to library management problems, postulate solutions and test these solutions in practice. The emphasis is on application and practical results - librarians are practical people and are much impressed by practical results. Such emphasis is likely to lead to the use of simple models rather than complex ones - to quote Cloot[3]:

> "If a management problem is complex and it is difficult to find a simple model for it, then it will be disastrous to try to find instead a complex one. With only a very few notable exceptions, successful models have been so simple that an operational research specialist would disown them".

(2) Lack of acquaintance with OR in the library profession.
Magic figures are dangerous in MI: if the manager does not understand the basis of calculation of figures presented to him he is likely to place either too much or too little reliance on them. It follows that if OR techniques are used in producing MI, the librarian must be acquainted with these techniques. It would be worthwhile:

---

1. LEIMKUHLER, F.F. op cit. p.18.

2. MORSE, P.M. Library effectiveness: a systems approach, MIT, 1968, p.vi.

3. CLOOT, P.L. Management information systems - can computers help? Computer Bulletin 11 (4) March 1968, 276-281. (Page 280).

(a) to run - primarily for professional librarians - short courses concerning the OR approach, OR techniques and their application in libraries.

It may be noted that although some 60 short courses are listed in the Library Association Record from January 1968 onwards, only two include in their titles the words "Systems Analysis" or "OR" or "Operations Research". In the two cases mentioned, the titles were "Systems Analysis"[1] and "Management in Libraries (particularly Systems Analysis)".[2]

(b) to produce a monograph in library systems analysis, including a section on OR and its (potential) application in libraries.

It may be noted that no reference to OR, systems analysis or mathematical models occurs in the indexes of a cross-section[3] of monographs on library management.

(c) for full-time library professional education curricula to deal more fully with OR and its (potential) application in libraries.

---

1. Library Association Record 70 (5) May 1968, p.129.

2. Library Association Record 71 (4) April 1969, p.122.

3. ASHWORTH, W. Handbook of special librarianship and information work. 3rd ed. ASLIB, 1967.

DOUGHERTY, R.M. and HEINRITZ, F.J. Scientific management of library operations. Scarecrow, 1966.

LOCK, R.N. Library administration. 2nd ed. 1965.

WHEELER, J.L. and GOLDHOR, H. Practical administration of public libraries. Harper and Row, 1962.

WILSON, L.R. and TAUBER, M.F. The university library. 2nd ed. Columbia University Press, 1956.

SOME DESIRABLE MANAGEMENT INFORMATION
FOR THE LIBRARIAN/MANAGER

Subject	Information
Staff	Organisation structure; job specifications; staff manual Notice of changes of employment, holidays; Reports from department heads.
Stock	Inventory of library stock; analysis of inventory as required, e.g. by subject, physical dimension, location, frequency of use, physical form; Account of acquisitions and losses with comparison of total shelf-footage used and available shelf-footage.
Buildings/ Equipment	Plan of buildings, including seating/shelving capacities; Schedule of equipment; Reports of periodic inspections.
Finance	Statement of all assets and liabilities; analysis of expenditure as required, e.g. all expenditure by expense code, stock expenditure by fund, subject, type of material; Regular statements of expenditure compared with budget.
Library Use data	Circulation: number of loans analysed by type of reader, subject, length of loan, type of material. Reader visits: number of visits analysed by type of reader, purpose of visit, number of tasks performed per visit. Enquiries: number of enquiries analysed by source organisation, medium (e.g. telephone/letter), subject, success in obtaining answer. Photocopies: number.

Subject	Information
Methods and Procedure	Procedures manual; unit costs of specified operations, e.g. cost of cataloguing the average book.
Other information	Inter-Library comparison comparative data for other selected libraries. Standards comparison of library data with standards. Inter-library loans/enquiries analysis of loans requested and enquiries passed to other bodies.

SOME SUGGESTIONS FOR BYPRODUCT MI TO BE PRODUCED BY COMPUTER-AIDED LIBRARY SYSTEMS[1]

System	Some possible Byproduct[1] Management Information	Use of information
Selection Data e.g. system for producing and distributing document references to individuals charged with selection. References might be obtained from MARC tapes	(1) Compare number of items and cost on MARC tapes with number of items selected and their cost; analysis of these statistics by subject, possibly also comparing budgets by (broad) subject.	Monitor selection by subject and adequacy of budgets.
	(2) Analyse selected items by source of suggestion and subject.	Monitor library staff/user contribution to selection process
	(3) Show mean and range delays between selection list and notification to library of selection by subject/fund. Possibly also by source of selection.	Monitor selection delays
	(4) Number and cost of selected but un-ordered items, cf. budget allocation and remaining budget allocation.	Monitor delay between notification of selection and order.
Order e.g. system for accepting document details (possibly for many items, from MARC tapes) and printing orders; maintaining an on order file.	(1) Number and cost of items on order cf. budget allocation and remaining budget allocation; analysis by subject/fund.	Monitor commitments by subject/fund.
	(2) Number and cost of items on invoices passed for payment; analysis by subject/fund and comparison with budget allocation and remaining budget allocation.	Monitor expenditure by subject/fund.
	(3) Average and range of item cost analysed by subject/fund.	Relevant to setting budget allocations.

1. Batch-processing computer systems are assumed since there are as yet very few on-line systems. The latter can be expected to yield considerably more information.

System	Some possible Byproduct Management Information	Use of Information
Cataloguing e.g. system for accepting catalogue data (say from MARC tapes) and for producing catalogues/catalogue entries	(1) Addition of dimensions of items by location; comparison of storage space of existing stock with total of available storage space.	Monitor amount and distribution of empty library shelf space.
	(2) The catalogue is the library's master record and provides data for much analysis of the library's stock, e.g. Number and age of stock by subject; the latter analysis could also be broken down by (a) accessibility, e.g. reading room stock/stack stock (b) availability for loan, e.g. loan v. reference only material.	Monitor library provision by subject
Circulation System in which each loan transaction for each loanable item is recorded and one or more files are maintained to show information such as material on loan; material on loan to each borrower; overdue material.	(1) Analysis of items loaned by subject, age of material, type of material; compare number of items loaned in each category with number of loanable items in that category.	Relevant to setting budget allocations by subject/department, to selection, and to choosing which material should go in more/less accessible locations.
	(2) Analysis of items loaned by number of times loaned and by type of borrower.	Relevant to consideration of loans policy, duplication.
	(3) Inter-library loans: analyse loans from other libraries by subject, age of material, type of material.	Relevant to selection and setting budgets.

DATA COLLECTION AND FEEDBACK

by

GEOFFREY FORD

June 1969

1. PURPOSES OF DATA COLLECTION

Data is important at four stages of library planning. In the first stage, it describes situations, in the second it tests theories; when plugged into suitable models, it explains situations; and finally, it predicts.

Data collection can be very expensive, so the reasons for collecting it must be made quite explicit before a finger is lifted to count a single piece of paper. Mere description in itself is of little use. A common statement in the calendars of universities is of the form "The library contains more than 250,000 volumes." This is peculiarly ineffective propaganda for the most part: it can convey little meaning to the average reader of calendars.

Description can be meaningful when used comparatively. Consider these three statements, all of them paraphrased from annual reports of university libraries.

> "The library added 10,000 volumes to stock this year."
>
> "Acquisitions this year totalled 10,000 volumes compared with 8,000 last year."
>
> "This year we acquired 6,000 volumes by purchase, compared with 10,000 last year; the fall was due to a rise in the prices of British books, the effects of devaluation, and a book-budget pegged at the same level as last year."

The first statement is meaningless, and the second is very little better. The third is packed with social, political and economic information.

At the hypothesis testing stage, it is sometimes necessary to obtain data from other sources than the individual library in question. What will be the effects of changing the borrowing regulations in a university library? On a simple level, lengthening the permitted loan period from two weeks to ten weeks seems likely to reduce the circulation of books. Is this so? Data from a university where this change occurred, unaccompanied with any other major change, is as follows.

	No. of students	No. of books borrowed	Books borrowed per head
Old regulations	4100	82000	20.0
New regulations	4600	77000	16.7
Percentage change	+12.2	-6.1	-16.5

The evidence does not refute the hypothesis and it is worth pursuing whether such a change will reduce the availability of books and thus reduce the quality of the service provided. If the change in regulations were accompanied by special provision of books in heavy demand, on the other hand, the quality of the service might increase. Is there evidence from elsewhere to support this? From a university which changed the loan regulations and increased the supply of heavily used books at the same time, we find the following.

	No. of students	No. of books borrowed	Books borrowed per head
Old regulations	2100	68000	32.4
New regulations		55000	23.9
Duplicate collection	2300	25000	10.9
Percentage change	+9.5	+17.6	+7.4

The effects of the two changes are difficult to disentangle. Taking the first two years of the new system, we find:

	No. of students	No. of books borrowed	Books borrowed per head
Year 1: New regulations		55000	23.9
Duplicate collection	2300	25000	10.9
Year 2: New regulations		55400	22.2
Duplicate collection	2500	35000	14.0
Percentage change	+8.7	+0.73 +40.0 } +13.0	- 7.1 +28.4 } +4.0

The evidence is still not conclusive, and not much is gained by extending the exercise. The need for a suitable model is demonstrated. The particular case of loan regulations has been covered in another paper (Buckland & Hindle, 1969) and this admirably illustrates the use of data in models. Model-building automatically leads to prediction. Prediction is a pre-requisite of planning. Librarians must plan, therefore they must predict. Of course, some predictions fail - that is why governments fall - but adequate information assists in the refinement of prediction. Some forecasts fail because not all factors can be controlled - weather, prices of periodicals, builders who go bankrupt; others fail because measures are taken as a result of the forecasts - Britain's export performance, extension of opening hours - and this kind of failure only pays tribute to the value of the exercise. Predictions based on a sound foundation of relevant data are more likely to be more correct than those based on hunch or guesswork.

If the population served by a library increases, what effects will this have on the use of the library? How can the library administration adjust to cater for differing demands? These and other questions can be answered if the right kind of data is collected in the right way.

2. CLASSES OF DATA

At the outset of problem solving, it is necessary to decide not so much how to collect the data, as what data to collect. Once it has been decided what is relevant, there is frequently an obvious way of obtaining the required information. The data concerning libraries can be grouped under four main headings: Resources, Activities, Operations and Background.

(a) Resources

These are usually the simplest of all to measure in terms of method, although not necessarily of effort. Labour, money, seats, shelves and bookstock are all resources which can be counted in obvious ways. This does not mean that they always should be counted. Since books occupy shelves, it is not necessary to count both, unless one wishes to check the average width of books in various subjects. In a library, one of whose tasks is preserving books for the future, it is largely irrelevant exactly how many books it actually possesses; it is necessary to know how many empty shelves there are, and how quickly they are being filled.

(b) Activities

Activities can be sub-divided into librarians' activities and users' activities, but this is to be avoided if possible because such a sub-division tends to overlook the impact of each upon the other. Thus the lending service can be measured in terms of how many books are borrowed in a day, or in a year, and deductions made about the requirements for issue-desk staff; but account must also be made of the users' time: an issue system which requires two to three minutes of a borrower's time for every book borrowed is that much worse than one which takes fifteen seconds.

So, the activities group includes the services of the library - lending, reference, photocopying, etc.; the "internal" activities - acquiring books, administration , etc.; and the users' activities - using the catalogue, consulting books in the library, and so on.

(c) Operations

Operations are the links between resources and activities. The activity "acquiring books" involves professional, non-professional and clerical labour, money, and shelf-space; "cataloguing" may use professional and clerical labour, and machine-time (for reproduction of catalogue entries). At the very least, a manager needs to know how much of each resource is absorbed by each activity. Just as it is inconceivable that a refrigerator manufacturer should be ignorant of how long it takes to make one deep-freeze, so it should be inconceivable that a librarian be ignorant of the amount of time required to file a hundred catalogue entries. Planning without this kind of data is less credible than science fiction of the Barbarella genre.

(d) Background

Background data is mainly concerned with potential demand and external factors that influence the actual demand. The size of the undergraduate body and the subject coverage in a university library; Reilly's Law, and the relative proportions of differing socio-economic groups in a public library. All librarians should be concerned in establishing why substantial numbers of people never use their library, with a view to providing the services which will reach these non-user groups.

3. METHODS OF DATA COLLECTION

Deciding what data to collect is largely a matter of the problem on hand. When planning a new extension to a library building it may be sufficient to calculate how fast the existing shelf space is filling up, in overall terms such as "3000 feet per year". When re-arranging material within an existing building it is necessary to be able to differentiate between subjects, as perhaps "Physics: 80 feet per year; Geology 43 feet per year".

(a) Problems

Each of the four classes of data has its own peculiar problems to be overcome with sufficient exercise of ingenuity and steadfast adherence to the principle of least effort.

(i) Resources

The measurement of resources such as staff, seats and money is automatic: the librarian knows how much money he has to spend, how many seats the raders may use, how many graduates there are on his staff. It may be necessary to classify the resources in different ways - professional staff, experienced and inexperienced non-professionals, clerical, etc.; seats for students, seats for professors, etc. - but that will depend on local conditions and is not a problem of measurement.

In such cases as bookstock there are opportunities for eliminating the painstaking volume by volume counting, as described later.

(ii) Activities

There is scope here for expending much effort for little return. The annual collection of statistics can degenerate into a meaningless ritual unless the object of the exercise is continually borne in mind. There is no point in knowing that only five per cent of inter-library loans are supplied through the regional bureau unless (for example) a decision is to be taken on the desirability or otherwise of maintaining a connection with it. Similarly, there is a danger in carrying out surveys of library use merely to satisfy curiosity. To know that one-fifth of all the users consult the catalogue on a particular day is not sufficient; it is necessary to know how this compares with other days, other times of year, other institutions; and if there are differences, what causes them?

Data collected in surveys of other libraries can often be useful. In the case of catalogue use, figures are available from a variety of sources.

	B'ham '64	Durham '66	Newcastle '68	UMIST '65
% of undergraduate users using catalogue on a given day	25	20	6	19
% of academic staff users using catalogue on a given day	Not available	26	15	19

Questions are instantly raised by comparisons of this sort.

The non-use of libraries occasionally gives librarians furiously to think: also the non-use of books by library users. Large numbers of students can always be found in university libraries before the summer term exams working solely with their own books. Data from a number of libraries suggests that the figure is around 50-60% of the total. The proportion is not inconsiderable at other times of year however.

All surveys in May	B'ham'64	Durham'68	Keele '67	Newcastle '68	UMIST '65
% of undergraduate users not using library materials	54	54	45	63	51

A curious aspect of non-use is provided by the category of users known as "External" (i.e. non-university users). It was noticed that at Durham in 1966, one-third of this category used no library materials when in the library; similar results were obtained in Newcastle in 1968, and another case has occurred in Sydney (Radford, 1967)!

(iii) Operations

This is possibly the largest terra incognita of librarianship. At any rate the literature is singularly deficient in information on the allocation of resources to activities. A recent study gives some data on the amount of labour allocated to particular tasks in certain types of academic libraries (Friedman & Jeffreys, 1969); but this is not related to the output resulting from the labour force. A recent work on productivity tells of a university library in which six professional librarians catalogued five books each per day (Butterworth, 1969): this suggests that one library at least believes that "each book is a law unto itself" (Berkowitz, 1961). This attitude appears to be quite common, and may prevent the collection of data on operations in all branches of library work. It may well be true that all books are exceptions (in which case why are there cataloguing rules?) It is certainly true that the time taken to issue a book should be independent of its intellectual content and physical characteristics.

In the absence of concrete data, some initial comparisons are possible at a very crude level. It is usually possible to find somewhere in Britain two libraries with similar characteristics. Data taken from the annual reports of two such libraries can throw up some interesting comparisons.

	No. of students	Academic staff	Library staff (Grad.)	Library staff (Others)	Loans per member of library staff	Additions to stock per member of library staff
University A	2850	420	21	24	3000	625
University B	3000	400	20	22	1800	350

Annual reports of other institutions can suggest many comparisons of this sort; a closer look at the reasons underlying differences can be of great assistance in exposing inefficient working or redundant practices.

(iv) Background

The problem here is of determining what is relevant. Other people's surveys can be of great use in highlighting particular aspects and suggesting new avenues of thought. Distance from place of residence to library (Luckham, 1967) - how does this operate in a university library situation? Is it obscured by the difference in need of different academic disciplines? Can librarians glean anything from the evidence garnered by behavioural psychologists, to assist them in planning the interior arrangement of libraries? There have been enough surveys of university libraries at least for the main pointers to be fairly clear. Closer looks at particular aspects are required now.

(b) Methods

The methods chosen for collecting data will depend on the precision required which will depend on the problem in hand. If it is really necessary to know exactly how many books on philosophy have been borrowed during the past twelve months, there is no alternative to counting the loan records one by one (unless an automated issue system is employed). In real life, where such precision is usually unnecessary, some short cuts can be taken.

(i) Getting machines to work: gates and burglar alarms

The simplest measure of the use of a library is the number of people visiting that library in a given period of time. The simplest method of counting them is by means of a meter mounted in a turnstile at the entrance or exit (although these sometimes upset the users). Daily recordings of this meter, accummulated over a period of time, readily show the weekly, termly and annual cycles of use, and can reveal trends which might otherwise not be apparent.

In some situations a recording turnstile is not feasible - perhaps for political, aesthetic or architectural reasons. It is possible to use a simple pressure-pad device - laid inside one quadrant of some revolving doors, or at some other restricted entry point - and this will be quite adequate.

A third possibility involves light beams and photo-electric cells; I have limited experience of these only, and that unhappy. It proved less reliable than either of the two previous methods for a variety of reasons, and was soon abandoned.

All counting devices have to be checked to see how accurately they represent reality; some characters always push a turnstile round twice instead of once, and so on. Spot checks made will show that readings are biassed while remaining consistent so that corrections can be made for the know bias.

In itself, the number of people using a library may seem of little account, but it does have a bearing on other activities.

The use of a computer as a means of data collection is covered elsewhere (Duchesne, 1969), and its possibilities are immediately apparent in the operation of such self-regulating loan systems as that described in another paper (Buckland & Hindle, 1969).

(ii) Avoiding tedium: weights and measures

Absolute precision in the counting of books, catalogue cards and loan records is almost invariably unnecessary. A tape measure and some letter scales are indispensable to the modern library manager.

When considering the reclassification of a section of the library, it is not necessary to know exactly how many books there are in that section; it is quicker to measure the number of feet of shelving occupied by the books and multiply by an appropriate factor. Thus, if we have 170 three-foot shelves occupied by books on German language and literature, we can say that we have about 5,300 books to reclassify (assuming a factor of 10.33 books per foot). This will be a sufficient basis for estimating the labour required for this activity.

A commoner exercise is counting the number of books catalogued in a week, or the number of cards filed in a catalogue. Again it is much easier to measure a pack of cards than to count them separately. The method is accurate to within 1%.

Naturally, this technique cannot always be used. Loan records must sometimes be counted, and where there are forms filled in by the users, or produced by a "Bookamatic" system, reasonable accuracy cannot be achieved with a tape-measure. The scales come into their own at this point, and again give estimates which are accurate to within 1%.

(iii) The democratic process: consulting the users

Some kinds of data can be collected sensibly in no other way than by means of a survey. In the context of an individual library, "survey" means a survey of the users. The general aspects of conducting surveys have been covered elsewhere (Line, 1967, and works referred to therein), and it is necessary here only to emphasise some points.

A. Instant diaries

"Instant diary" describes a particular kind of survey by questionnaire. Each user, on entering the library, is handed a form which he fills in during his stay, returning it to the survey staff on leaving.

It is essential that the form used by brief and easy to complete.

A user who has just dropped in to return a book does not want to sit down for five minutes or more to answer a detailed questionnaire. In Durham, cards have always been used because they can easily be filled in while standing up. The best size is a postcard: some identification details on one side, two or three simple questions on the other - the survey staff can stamp on the times of entry and exit if these are required. Such a card wastes as little of a user's time as possible (20 seconds if he is awake) and therefore induces low "consumer resistance". Of course, one can never overcome the problem of the professor who will take five minutes to explain why it is a waste of his time to spend thirty seconds filling in a card; but an undergraduate revising for Finals will spare thirty seconds where he feels that five minutes is an imposition.

The instant diary survey is of great value in obtaining information on a variety of the users' activities and is a relatively cheap way of obtaining such information. Appendix 2 gives a description of a survey of this nature carried out in May of 1968.

B. Questionnaires

"Questionnaires" here means the kind of survey in which forms are disseminated (usually by post) to the members of a group, or selected individuals of that group, with a request that they complete and return the forms in thier own time. This kind of survey can be used when it is desired to assess the relationship of the library to total information activity; or to assess the place of the library in the lives of the population generally rather than in the lives of those who happen to use it on a particular day. The questionnaire is a less precise method of gathering detailed data about library use - questions must be general, memories fade, and facts are more easily obscured by opinions. For example, an instant diary survey was carried out in Durham for a week, and the average length of time spent in the library by undergraduates was established. The following week, a questionnaire asked them how long they had spent in the libraries in the week of the survey. On average, they over-estimated by 60%. This was probably due to a combination of faded memories and an assessment of their personal costs in using the library.

There is a tendency to make questionnaires too long and complicated. Perfection is difficult here: some of our own took far too long to fill in, even if the respondent knew all the answers. Twenty minutes is the absolute maximum one can expect, unless the rewards of giving the information are immediately apparent to the respondent.

For example, one question asked was of the form "At what times of day did you use the library last week?" Three periods of the day were given (9-1, 1-5, 5-9) for each day of the week.

It is extraordinarily difficult to remember this kind of information, short of keeping a diary, and there is not much incentive in forcing oneself to remember. It is very easy to spoil the sample by holding too frequent surveys or in trying to collect data which does not seem to have any conceivable relevance other than satisfying mere curiosity.

A useful technique is that of starting off with a general questionnaire, followed not less than six months later, by another which deals in detail with specific points thrown up by the previous one. One early questionnaire in Durham asked the academic staff there to describe their main sources of information. On the basis of their replies, a fuller questionnaire was designed asking them to indicate the relative importance of certain specific sources.

C. Interviews

I always read with astonishment the accounts of surveys in which users are stopped on leaving the library and questionned as to their motives, actions, attitudes and so on. I have always felt that this is too much of an imposition on their time, particularly in a university where the users may be undergraduates rushing to lectures, lecturers rushing to committee meetings, or research students rushing to catch grasshoppers or pyridine thianamide. The only ones with time to talk are those who insist that surveys are a waste of time. This kind of survey is almost inevitably selective, unless there are vast teams of interviewers available, and it is usually not possible to employ students for this kind of work, unlike the instant diary survey (see Appendix 2).

Interviews can be of use however in some situations. If it is thought that a postal questionnaire would have poor response; if a very detailed study of information requirements and use is under way; if a particular aspect of library services is being evaluated; or if some pointers are needed as to the kind of library services to provide in the future. A small scale interview survey can be used as a means of highlighting particular aspects which can then be covered in a wider survey by questionnaire.

As with questionnaires, the questions must be very carefully composed. I was involved once in a survey in Harrogate. We were trying to establish the basic knowledge of libraries which students at a college of further education possessed at the beginning of their course. This data would then be of use to the librarian in planning his courses of instruction for these students. The interview schedule was designed with 15 to 18 year olds in mind; we drew a random sample; and wheeled them in. My first interviewee, unimaginably, was in her sixties and half the questions succeeded in being irrelevant and impertinent.

D. Observation

Whenever possible library use should be surveyed by direct observation. Catalogue use, seat occupancy, queue formation are all obvious candidates, and some data on operations can be assembled in this way also. It is very easy to time how long an individual user has to wait at the issue desk when he borrows or returns a book: how long he spends in the queue, and how long waiting while the issuing or discharging of the book takes place. The behavioural scientist is interested in studying the working habits of users - reading, writing, etc. One cannot follow the users around all of the time - although one of our temporary assistants tried to do so once (he said it was a fault in his briefing); but some kind of activity sampling is possible. In a large reading-room, from a suitable vantage point, it is possible to observe, at pre-determined times, what is apparently being done by the occupants of pre-selected seats; or the use made of a selected area of the library can be studied - how many users wandered along the shelves, how many took books away, and so on.

Before doing any of this, the reasons for doing it must be absolutely clear. It does not help anyone much to know that readers spend 20% of their time gazing into space; no librarian would put up a notice saying "KEEP YOUR EYES ON YOUR BOOKS" - though I did see a notice in a library once which said "PLEASE CLEAR YOUR BOOKS OFF THE TABLES".

4. USING THE DATA

It is not necessary to be a statistician in order either to collect data or to use it. There are many simple ways of using data in libraries; and series of data built up as a result can also be used in the more sophisticated techniques of operations research. The immediate relevance of the data is important, particularly if a survey is conducted. The user is happier if he can see some direct improvement in service as a result of being subjected to a questionnaire. The following case-studies relate to university libraries.

(a) Case study 1: short-term allocation of resources

This illustrates what might be called "micro-planning" - making sensible decisions about allocation of staff duties from week to week. Apart from the routine day to day work of the library, there are always less urgent tasks to be fitted in somewhere. It might be that shelf-tidying is to be done every week. When should it be done?

Now this library has a turnstile, and daily readings have been taken for the past few years. From these readings it is immediately apparent that there is a weekly cycle of use -

the busiest full working day is Monday, the least busy is Friday.

	Average Turnstile Reading Monday	Average Turnstile Reading Friday	Ratio Monday/Friday
Oct.-Dec. 1966	453	334	1.36
Jan.-Mar. 1967	543	420	1.29

It is worth trying to discover whether the activities of the desk assistants are directly affected by the number of users. So, in one week, a record is kept of the work done by the desk staff.

	Monday	Friday	Ratio Monday/Friday
Turnstile reading	620	480	1.29
Books shelved	474	407	1.16
Filing and unfiling of loan records	406	296	1.37
Total "item-actions"*	1015	746	1.36
Minutes of staff time available	1080	1080	1.00

2 assistants are available

* item-action: one action concerning one item (e.g. one book shelved, one loan-record filed).

The overall results support the hypothesis, and suggest that one assistant could look after the issue desk on Friday afternoons and thus release the other for shelf-tidying unless, of course, all the Friday users come in during the afternoon. This is unlikely, however, as a survey of library use showed that, at a "normal" time, twice as many users entered the library in a morning as in an afternoon.

(b) Case study 2: estimating requirements in the medium term

By extrapolation from the previous study it is possible to lay out some requirements for planning on a year to year basis.

Some aspects of library staff activity have been shown to be linearly dependent on the amount of use made of the library. If some sensible predictions can be made about the future use of the library, it is then possible to estimate the amount of junior staff required. If this can be combined with predictions concerning stock acquisition, some light will have been shed on the darkness ahead. Can such predictions in fact be made?

Some crude measures of library use are already available -

daily turnstile readings, number of books borrowed during the year; since the first of these has implications for the library staff activity, so has the second. What can be used as a base-line for predicting the future levels of these activities?

Universities exist because there are students, so that the number of students should affect the use of a university library. Given that

$$\text{Library Use} \propto \text{Number of students}$$

then

$$\frac{\text{Number of students next year}}{\text{Number of students this year}} = \frac{\text{Library use next year}}{\text{Library use this year}}$$

The table shows the actual figures and the predictions based on the above relationship.

	1966/67 Actual	1967/68 Predictions	1967/68 Actual	Error in Predictions
No. of students	2810	-	2990	-
No. of loans	12960	13790	13980	1.4%
Max. daily reading of turnstile	1095	1165	1180	1.3%
Max. weekly reading of turnstile	5790	6160	6010	0.8%

At first sight, it looks as if the proposition is proved. However, a closer look is needed. What if the error in the prediction should be larger than the actual increase in the figures concerned?

	No. of students	No. of loans	Max. Daily Turnstile Reading	Max. Weekly Turnstile Reading
Percentage increase 1967/8 over 1966/7	6.4	7.9	7.8	3.8

On this basis we can say that the actual increases in the levels of activity are five or six times as great as the errors in the predictions. As long as this holds true, extrapolation of this kind may be useful on a year to year basis. This is not to say that student numbers hold good as an indicator in all universities; and in any one institution, a change in regulations or services can have a considerable effect on the levels of activity.

A point of wider interst is raised by the second table. The number of loans is increasing more rapidly than the number of students. This suggest that there may be a change occurring in the reading habits of students. Why is this? - The data has suggested an important problem which is of great educational significance, and of interest to a wider audience. So this trend must be investigated - there may be strong opinions within the university as to whether this is a good thing or not.

Another point which is raised from perusing the data. The maximum daily reading of the turnstile has increased - twice as much as the maximum weekly reading. This suggests that the use of the library is becoming more uneven over the week, and this may cause some timetabling problems (for the library staff).

The other half of the problem is predicting the rate of acquisition of books. A hypothesis worth testing is whether the published data on average book prices can be used as a guide. In fact, the sample we tried first in Durham had an average price of £2.15. 0. - almost exactly the average price for adult non-fiction books in B.N.B. that year (B.N.B., 1968). This was so exciting that we pursued it further - into the realms of disappointment. The average price did not, as we had hoped, fluctuate in the same manner as the national figures - as why should it? since universities buy foreign books as well. Not only did the local figures vary widely from the national ones, they also decreased in some years when the national figures rose. Some quick calculations on published data from other universities also showed an erratic pattern. This suggests that maintaining a fixed number of cataloguers may be an inefficient method of working, as they are likely to find time on their hands in one year and to be overworked in another. A more flexible staff structure seems necessary, so that in the lean years the librarians can promote the use of the literature, and in the fat years they can concentrate on the provision. At all times a continuing appraisal is required of the use of the literature actually purchased. This is particularly true of periodicals, with their heavy continuing commitment. Some typical findings can be usefully summarised here.

(i) A sample was drawn from the accessions register of books added to the library. After these books (originally suggested for purchase by the academic staff) had been on the open shelves for one year, they were examined for evidence of use. 45% had never been borrowed.

(ii) Users were asked to initial the issues of current periodicals which they consulted. Preliminary results indicated that 20% of the titles had never been consulted; there were also reports that some users initialled periodicals which they did not themselves use but which they thought should be kept as insurance.

(iii) A collection of duplicate textbooks was provided in a university library. All the books included in it were nominated by the academic staff or suggested by students as essential reading. After two full terms of availability over 40% of the books had not been borrowed. This is not to suggest that either academic staff or students are at fault - a complete year's data not being available.

(iv) A discussion occurs on the relative place of the university library and departmental libraries. Books for purchase are selected by the academic staff for all libraries, and it is stated that "Infrequency of use is one of the main criteria for selecting books for the university library".

The argument "We are building for posterity" may be valid, but continuing this process to extremes has ludicrous results. It may well be that more would be gained by diverting some funds to purchasing alternative sources of information, and there is some data to support this.

(c) Case study 3: medium-term allocation of resources

Another paper at this seminar deals with the application of linear programming to library planning (Morley, 1969). This sophisticated technique, using a few seconds of computer time, yields much information on possible future policies. Data on resources, activities and operations is required - all of which can be obtained relatively easily. The description here refers to the data collection necessary for the calculations relating to the Durham Arts/Social Sciences library during one term.

(i) Resources

It is important to work within a particular model, in order to decide which resources are relevant. This will be determined by the activity considered, and the operations involved in those activities. If it is decided to exclude Administration and Committee work from the model, then the resources devoted to these activities must be deducted from the total available.

In a simplified model, the bookstock of the library is not taken into account. The shelf-space available for new books is relevant, and also the amount of shelf-space occupied by one volume. An accurate estimate of the former was available in Durham; the latter was not known with any precision, and this had to be found. Some 300 well-filled shelves in all parts of the library were selected; the books on them were counted, and the amount of shelving occupied was measured. As a result, ten books per linear foot was adopted as a standard measure - this accorded well with data from Newcastle University which later became available.

(ii) Activities

Information as to the activities comes from various sources. The number of loans made is a standard compilation, as is the number of inter-library loans obtained and provided. The number of books added to stock is derived from weekly figures kept by the cataloguing department. Other figures are estimates based on surveys. Turnstile (or rather, pressure pad) readings are available covering the whole period under consideration; from spot-checks made previously, the bias of the recording device is known, and an accurate estimate can be made of the numbers using the library. A survey in May of 1968 provided a check on the likely numbers of users who worked in the library without library materials and the average time spent in the library. Earlier surveys (in 1966 and 1967) had shown approximately what proportion of users had consulted the library staff, and checks on this were provided by data from Newcastle (1968) and Edinburgh (1967). From these surveys it was also known how long was spent in the library by users who had only consulted the library staff during their visit. This evidence suggested that on the whole, undergraduates make "short" consultations, and academic staff make "long" consultations. A reasonable assumption to make is that undergraduates consult junior library staff (on the issue desk), and that academic staff consult senior library staff. On this basis the activity levels for user services by senior and junior librarians were calculated.

(iii) Operations

All of the figures concerned with operations had to be calculated from first principles, in order to provide the technical coefficients for the linear programming calculation. To do this completely would require an extended method study and a lot of effort. Some short cuts can be taken, however, which make the work manageable.

The activity described as "Increasing library stock" (I) covers a wide range of operations. Initially, it involves senior librarians' time - checking and approving a recommendation for purchase; clerical time - typing an order; junior librarian time - filing the order slip. When the book arrives it may be in a parcel - unpacked by porters; it is then accessioned - junior librarian; an invoice will be checked - senior librarian. Then the book is catalogued and classified - senior librarian;

processed - junior librarian; and shelved - porter, junior librarian. The catalogue entry is typed - clerical; checked - senior librarian; reproduced - clerical; and eventually, filed - senior librarian. Not many of these activities are amenable to timing for individual titles, for obvious reasons. Over a period, it is possible to make some estimates which will be close to the truth. If a team of four cataloguers deal with 270 titles in a week, that is about 32 minutes per title ("national average" cf. Friedman & Jeffreys, 1967). If the assistant librarian (acquisitions) can check 40 recommendations in an hour, and pass an invoice for 20 items in 40 minutes, that is a total of 3.5 minutes of his time per title. In this way the technical coefficients for the activity "Increasing library stock" are gradually built up. A point to note is that the user's only concern with this particular activity is the book on the shelf, and the means of locating it. The rest is purely librarianship.

A similar process was necessary for the other activities. "Working with library materials in the library" (LML) is one case requiring estimates derived from surveys. The only library staff time involved is that needed for shelving books left on the tables by the users. The survey in May 1968 suggested that two-thirds of the books were re-shelved by the users themselves. As a result, it was calculated that every two hours spent by users in the library working with library materials resulted in one book requiring to be shelved by the library staff.

Another activity requiring a variety of separate operations was "Long loans" (LL). This covered the operations of issuing the book, discharging it on return, and shelving, with the associated filing activities. Some trials were made - such as timing the filing of 100 loan records, and checking on how long it took to discharge a loan. The latter was extremely variable - anything from 27 to 91 seconds, depending on how many books on a particular subject were on loan, and how many books the borrower had on loan altogether.

A final check on the overall accuracy of the data is possible.

	I	OILL	LML	LL	SL	Resource Available
S	12.0	0.7	0.0	10.0	0.0	2004
J	3.0	12.0	6.0	80.0	50.0	2330
Activity level	153	24.8	53.7	15	1	

Obviously, the total amount of a resource used must be not more than the total available. So, if each activity level is multiplied by the amount of a resource required for that activity, and the products are added, the total must be less than or equal to the total resource available. Thus from the table above

(153x12)+(24.8x0.7)+(53.7x0)+(15x10)+(1x0) must be not greater than 2004;

and (153x3)+(24.8x12)+(53.7x6)+(15x80)+(1x50) must be not greater than 2330. Both of these conditions are satisfied, and so our data is on the whole correct.

When considering the fugure, it may be necessary to take account of new techniques (which will affect the technical coefficients), and increasing demands (which will affect the activity levels). It may be that the resources will change also. Predicting future demand has already been touched upon; and different technical coefficients can often be imputed from other libraries already using the techniques which are being considered for adoption.

(d) Long-term planning

For long-term planning it is necessary to know far more details of the effects of particular factors on library use. Some of the data required may be available from surveys made in other localities. A major consideration in the siting of a new library will be the distance of the place of residence of the users from the library. A survey at Edinburgh in 1967 showed that the further away from the library a department was situated, the less likely were the academic staff to visit the library. This is in part supported by surveys in Durham in November 1966 and in March 1969.

November 1966: Graduate students & academic staff

Arts Depts. Distance from library	<200 yds	200-500 yds	500-1000 yds
No. of visits/member of dept.	1.38	1.23	0.94

March 1969: Undergraduates

Halls of Residence Distance from library	<100 yds	<400 yds	>1200 yds
% of potential visitors who used the library after 6 p.m.	48.5	40.5	23.6

Another consideration is the influence of social custom, or the "group norm." Again from the Durham 1966 survey we find that on the whole, the distance of the place of residence from the library had no influence on the use made of the library by undergraduates except for those who lived within 100 yards (this is not necessarily contradictory of the table above, which relates only to evening use of the library).

In cases where a restricted site is available for a new library, it may be thought necessary to divide the library collections. Evidence on the desirability or otherwise of this step can be gained by a study of loan records: an automated issue system readily yields information on who uses what books. This data is tedious to collect by ordinary manual means, but it can be very important as an indicator of the degree of inconvenience caused by splitting library collections. A pilot study of the loan records at Newcastle University in 1967 showed that of the books borrowed by members of staff of the Faculty of Agriculture, 26% were classified as Agriculture, 26% as Arts and Social Sciences, and 47% as Biology and Medicine. In Sheffield in 1960-61 the same theme occurred (Saunders et al 1966), as also at Leeds (Page and Tucker, 1959), and Birmingham (Humphreys, 1964).

A more complete approach to long-term university library planning involves not only considerations of the number of books which are likely to be acquired in the next forty years, but also of new library technology; the working habits of students and academic staff; the influence of teaching methods on library use, and other aspects of educational technology as yet untouched by research.

5. FINAL SECTION

The more sophisticated techniques of operational research frequently require series of data which are not normally collected. Library management is already complex and is likely to become more so, but the use of advanced techniques will help to simplify it again, provided the data is available. The fact that such data can already be used in simple ways to ease the burden of decision-making may encourage librarians to make provision for the more advanced techniques being developed. Sufficient surveys of the use of university libraries have been carried out for some general stateements to be made concerning such use. The need now is for specific investigations on the availability of books (such as those made at Lancaster); on the influence of in-library factors on book-use (see for example, Harris, 1967); and other explorations into the factors affecting library use.

One dictum is worth repeating: there is a strong "time-preference" in data collection - i.e., the quicker the pay-off, the happier are both those who collect it and those who have been bothered with surveys.

Acknowledgments

The library users and library staffs at Durham and Newcastle Universities for their forbearance.

Miss A. M. McAulay and Mr. W. B. Woodward of Durham University Library for making available data for me to use.

Richard Morley for reading and criticising this paper.

The PEBUL team, especially Bill Morris, survey organiser par excellence, whose work provided the opportunities for learning about data collection.

References

B.N.B. 1968 "Average book prices." *Libr.Ass. Rec.*, 70: 233-238.

Berkowitz, A.M. 1961 *A study of the costs of cataloguing books with Library of Congress catalog cards and by original cataloguing methods*. Catholic University of America. (M.A. Thesis).

Buckland, M.K. & Hindle, A. 1969 *Loan policies, availability and duplication*. Lancaster. (Paper presented at the Seminar on Planning Library Services).

Butterworth, J. 1969 *Productivity now*. Oxford, Pergamon Press. 148 p.

Duchesne, R.M. 1969 *Management information from computer aided systems*. Lancaster. (Paper given at the Seminar on Planning Library Services).

Friedman, J.E. & Jeffreys, A.E. 1967 *Cataloguing and classification in British university libraries: a survey of practices and procedures*. University of Sheffield, Postgraduate School of Librarianship. 37 p.

Friedman, J.E. & Jeffreys, A.E. 1969 "Cataloguing and classification in British university libraries. Part 2: the Labour force." *J. Docum*, 25: 43-51.

Humphreys, K. W. 1964 "Survey of borrowing from the Main Library, the University of Birmingham." *Libri*, 14: 126-135.

Line, M.B. 1967 *Library surveys: an introduction* ... London, Clive Bingley. 146 p.

Luckham, B.L.G. 1967 "The distribution of public library members in Southampton by distance of residence from library premises." *Res.in Librarianship*, 1: 169-170.

Morley, R. 1969 *Project for Evaluating the Benefits from University Libraries*. Lancaster. (Paper presented at the Seminar on Planning Library Services).

Page, B.S. & Tucker, P.E. 1959 "The Nuffield pilot survey of library use in the University of Leeds." *J. Docum*, 15:1-11.

Radford, N.A. 1967 "Intra-mural use of a university library." *Aust. Libr. J.*, 16: 209-216.

Saunders, W.L. *et al* 1966 "Survey of borrowing from the University of Sheffield library during one academic year." In *The provision and use of library and documentation services* (ed. by W. L. Saunders). Oxford, Pergamon Press. 198 p.

Additional reference

Harris, I.W. 1967 *The influence of accessibility on academic library use*. Rutges - the State University, New Brunswick, N.J. (Ph.D. Thesis).

APPENDIX 1

Surveys carried out by PEBUL, November 1966 - May 1968, in the Universities of Durham and Newcastle

1. Instant Diary Surveys

Code No.

A1D	All users of the Durham Arts/Social Science Library: 17-23 Nov. '66.
A2D	All users of the Durham Science Library: 21-26 Nov. '66.
A3D	All users of the Durham Oriental Library: 21-26 Nov. '66.

These three surveys used the same card form, to elicit information on the extent and nature of library use. Response was ca 96%. Materials cost £10. Extra labour £45. The card took 90 seconds to complete.

The same form was used again in one-day surveys six months later.

A12D	Arts/Social Sciences:	16 May '67.
A22D	Science:	11 May '67.
A32D	Oriental:	18 May '67.

No extra labour was employed. Materials cost was included in that for the first group.

F1D	All users, Durham Arts/Social Sciences Library: 23 Nov. '67.
F2D	All users, Durham Science Library: 28 Nov. '67.

Response: ca 96% Materials cost: £7.
Extra labour: nil.

Postcard size forms were used; information was sought on the areas of the library used, and on the subjects of the books used inside the library. An attempt was also made to discover whence library users had come, and whither they were going. The card took 60 seconds to complete.

K1N All users, Newcastle Main Library: 22 & 27 Feb. '68.

Response: 97% Materials: £38 Extra Labour: £13

A card form was used. Information was sought on extent and nature of use in the library; areas of the library used; subjects of books consulted; place of origin and destination of library users. The card took about 2 minutes to fill in. The card combined features of those used in the A series and F series of surveys (above) and likewise was partly easy to analyse and partly difficult.

The same card was used two weeks later.

K4N All users of 8 departmental libraries, Newcastle: 12 Mar. '68.

Response: ca 96% Materials: see K1N Extra labour: £10

M1D All users, Durham Arts/Social Science Library: 6 May '68

M1N All users, Newcastle Main Library: 6 May '68

M2D All users, Durham Science Library: 6 May '68

Response: ca 96% Materials: £17 Extra Labour: £12

A postcard was used asking for information on the number of books used in the library. The card could be completed in 20 seconds. Analysis was simple.

2. Questionnaire Surveys

B6D Durham Academic Staff. Response: 82%

B8D Durham Research Students. Response: 56%

Materials: £3 Distributed mid-Nov. '66

A single foolscap sheet, most of which was left blank for the recipients to describe their information seeking habits, and to let off steam generally. The data thus gathered was used to design the form for a subsequent survey (D6D).

Analysis was more interesting than simple.

C9D Durham undergraduates. Response: 72%

Materials: £8 Extra Labour: £95 Distributed late Nov. '66

Two single-sided foolscap sheets, seeking information on library use and working habits of undergraduates.

Straightforward tabulation of replies is easy; correlation difficult because of the sheer bulk of the data.

D6D Durham Academic Staff. Response: 73%

Materials: £2 Distributed mid April '67

A single foolscap sheet, asking recipients to rank named information sources (derived from their replies to B6D). Also asked about vacation use of libraries outside Durham.

Analysis simple.

G9N Newcastle Undergraduates (20% stratified sample).

Response: 61% Materials: £2 Distributed late Jan. '68

Five foolscap pages. Most questions fairly easily answered. The survey was concerned with library use and working habits, and attitudes to university and departmental libraries.

Analysis quite difficult for some questions.

H7N Newcastle Junior Academic Staff Response: 35%

H8N Newcastle Postgraduate Students Response: 79%

Materials: £24 Distributed late Jan. '68 -
mid Mar. '68

Six foolscap pages. Questions on library use, information seeking habits, working habits. Some questions required thinking about. A minimum of 20 minutes required to complete this questionnaire.

Analysis quite difficult for this type of questionnaire.

J6N Newcastle Academic Staff (down to lecturer)
Response: 58%

Materials: £14 Distributed mid Feb. '68

Four foolscap pages. Questions on library use and information seeking habits. Also asked to rank information sources (based on D6D above). As in H series, questions often required thought.

Analysis again quite difficult.

L6D Durham Academic Staff who receive Social Science Current Awareness Service Response: ca 95%

Materials: £1 March '68

A preliminary evaluation of the current awareness service. Questions required some thought, but assistance was given where necessary by one of the Project leaders.

Analysis simple.

---

3. Survey Ratings

Series letter	User Inconvenience	Difficulty of Analysis
A	**	**
B	**	**
C	**	***
D	**	*
F	**	***
G	***	*****
H	***	*****
J	***	*****
K	**	***
L	**	**
M	*	*

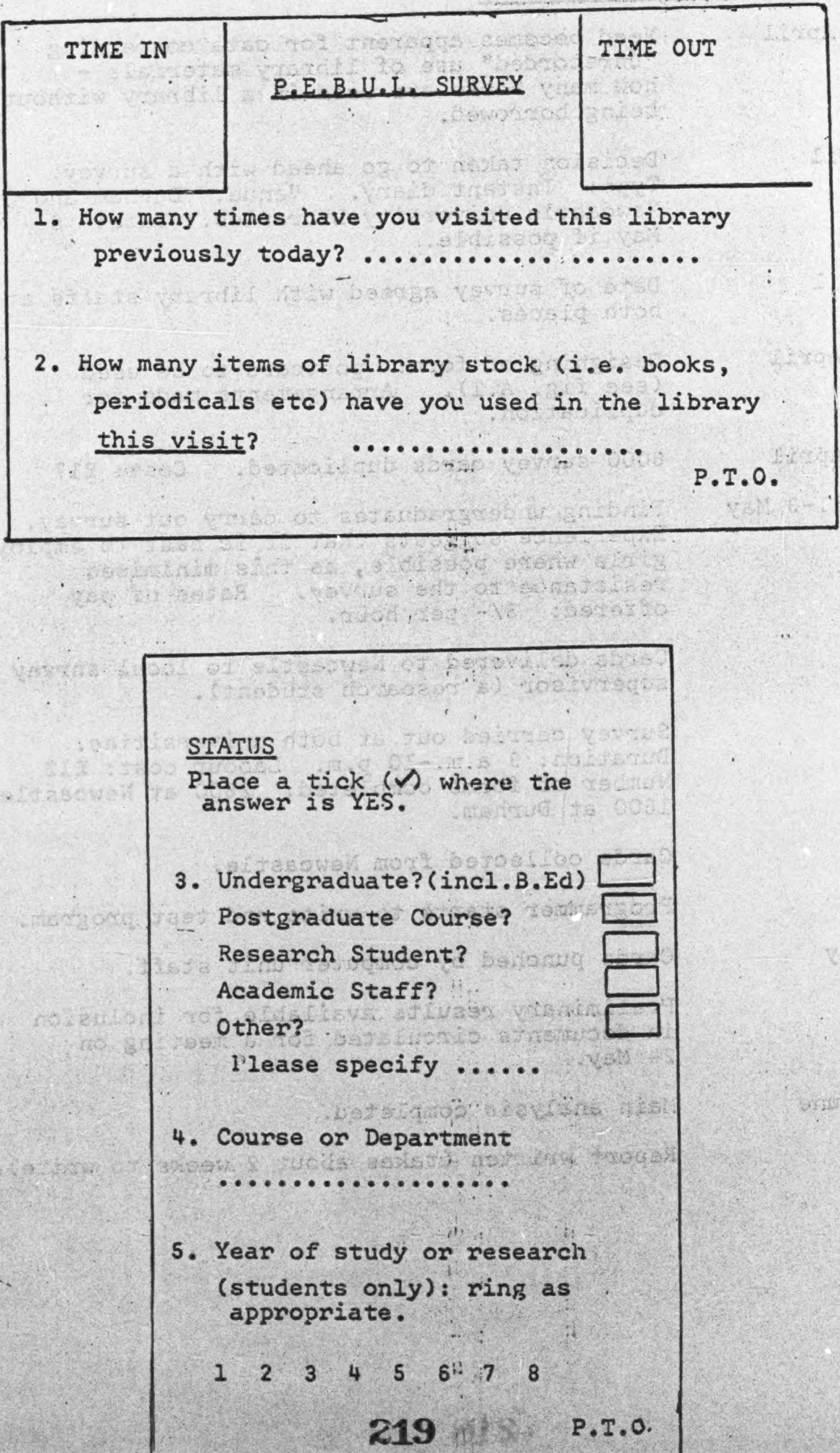
The fewer the stars, the better the survey.

APPENDIX 2

An outline history of a survey

Early April	Need becomes apparent for data concerning "unrecorded" use of library materials - how many books are used in a library without being borrowed.
18 April	Decision taken to go ahead with a survey. Type: Instant diary. Venue: Durham and Newcastle University Libraries. Date: 6 May if possible.
19 April	Date of survey agreed with library staffs at both places.
19-26 April	Designing of form: postcard to be used (see fig. A.1). Arrangements made for duplication.
26-30 April	6000 survey cards duplicated. Cost: £17
26 April-3 May	Finding undergraduates to carry out survey. Experience suggests that it is best to employ girls where possible, as this minimises resistance to the survey. Rates of pay offered: 5/- per hour.
3 May	Cards delivered to Newcastle to local survey supervisor (a research student).
6 May	Survey carried out at both universities. Duration: 9 a.m.-10 p.m. Labour cost: £12 Number of forms completed: 2800 at Newcastle, 1600 at Durham.
7 May	Cards collected from Newcastle.
8 May	Programmer starts to write and test program.
8-10 May	Cards punched by computer unit staff.
15 May	Preliminary results available for inclusion in documents circulated for a meeting on 24 May.
Early June	Main analysis completed.
Later	Report written (takes about 2 weeks to write).

Fig. A.1



TIME IN

TIME OUT

P.E.B.U.L. SURVEY

1. How many times have you visited this library previously today? ....

2. How many items of library stock (i.e. books, periodicals etc) have you used in the library this visit? ....

P.T.O.

STATUS

Place a tick (✓) where the answer is YES.

3. Undergraduate?(incl.B.Ed) ☐
Postgraduate Course? ☐
Research Student? ☐
Academic Staff? ☐
Other? ☐
Please specify ......

4. Course or Department
....

5. Year of study or research (students only): ring as appropriate.

1 2 3 4 5 6 7 8

P.T.O.

Models and Measures for non-profit making services

By: A. Hindle
Department of Operational Research
University of Lancaster

I. INTRODUCTION

This paper considers the role of Operational Research in social planning. Since this is a relatively new field of application it is important that the rationale of the Operational Research approach is critically reviewed. By examing work in this field I therefore aim to highlight the major difficulties encountered.

Planning can be defined as anticipatory decision making. It is an inherently difficult activity in that it usually involves a complex set of interacting decisions and is concerned with future activities in a dynamic environment. Wrong decisions will usually prove costly because of the fact that plans ordinarily involve the provision of expensive facilities having long lives and inflexible uses. Alternatively, but equally critically, the plans may concern sweeping reforms which will be difficult to reverse.

The special difficulties of applying the Operational Research approach to social planning derive from the nature and organisation of community services. These difficulties are of two types. Firstly models of the planned process may be difficult to construct especially if the success achieved by the plan depends significantly on community behaviour. Secondly, measures of effectiveness are usually very difficult to devise, not only because of the usual problems of combining objectives measured on different scales but also because the appropriate value system is often not easy to identify.

II. MODELLING THE PLANNED PROCESS

Planning is a creative activity concerned with preparing the future: of fundemental importance is the definition of that 'desirable' state towards which the organisation should be heading: the goals and objectives of the organisation. However part of all planning consists of the selection from amongst alternative plans and in order to do this it is necessary to predict the performance of each planned process. The costs of implementing each plan can then be compared with its expected benefits and the best plan identified.

In this paper, I am, of course, mainly concerned with the use of explicitly formulated scientific models for predicting performance. We should, however, be unwise not to recognise that science has at present a limited capability to construct explicit models and should not, therefore, underestimate the human capability to construct implicit ones. Of course, in many situations the predictive power of scientific models has proved far superior to models derived in other ways, but in relation to some phenomena science is still making a somewhat hesitant start. This is particularly true of the explanation and prediction of human behaviour. Very often in fields such as health and education the relevant aspects system performance depend significantly on human behaviour and these deficiencies in basic scientific knowledge therefore prove a stumbling block in social planning.

For the purpose of subsequent discussion it is convenient to subdivide the types of models relevant to planning into the following three categories:

(a) Models for forecasting the changes which will occur in the environment if no plan is implemented.

(b) Models for predicting the performance of plans in a hypothetically fixed environment.

(c) Models for describing the way in which the environment will respond to the planned process or be affected by it.

In social planning these often correspond to models of demand for community services, the capability of the planned process to meet a given demand and the way the process will affect the demand. In order to clarify the nature of these three types of models it may be helpful to discuss a problem of planning in the police field outlined by Sargeaunt (1966). The first part of this problem is to predict the increase in number of crimes committed if no significant changes are made in police activity. The second part of the modelling process requires an estimate to be made of the way in which plans to increase manpower, improve forensic methods and introduce superior equipment will increase crime prevention and detection. At the final stage it is necessary to model the effect of the measures designed to increase crime prevention and detection on criminal activity.

Before discussing these model types in more detail it is important to note that outputs required from models for community services vary a great deal in general nature and are in some cases difficult to quantify.

Models for forecasting changes in the environment are usually very unsophisticated. Simple assumptions are frequently made about the form of the model and then its parameters are estimated from historical data. For example, in hospital planning forecasts of morbidity rates are usually obtained from straight-line extrapolations of past statistics. Forecasts of this type are notoriously unreliable not least because of the possibility of radical technological developments. Thus prediction of the requirement for tuberculosis sanatoria derived in this way following the last war resulted in the building of many now useless buildings in remote country areas. An alternative to basing forecasts on past data is to rely more on relevant expert opinion. This approach has been used in a very detailed way in planning urban renewal. Surveys have been undertaken, for example, in which qualified building inspectors estimate the useful lives of existing properties. A case in which these two forecasting approaches were combined is illustrated by the government white paper, "The Long-Term Demand for Scientific Manpower" (1962). In this, estimates of demand are obtained by making assumptions about the permanence of critical ratios such as work force to population, manufacturing industry to other industry, students to teachers, etc. and then projecting trends in population growth, industrial development, and teacher supply twenty years into the future. The forecasts derived were then doctored to allow expert opinion to modify the outcome.

Beer (1962) pinpoints the precarious nature of these forecasts by noting that if, as is possible, the level of technological sophistication in the iron and steel industry was raised to that already existing in the oil, chemical and electronics industries the increased demand produced would of itself, be greater than the total increase projected. As a result of the unreliability of these forecasts Beer concluded that it is better not to attempt detailed forecasting of an uncertain future but that it is preferable to design processes which are adaptable - "Away, then, with attempts to forecast the state of affairs in 1970 - which is known only to God." Ackoff (1966) takes a very similar line in considering that planning should be concerned with making provision for the inevitable.

Modelling the performance of a planned process in a given environment has proved easiest where significant phenomena are physical. Thus, although details of work in the military field are not fully published it seems that appreciable successes have been achieved in modelling weapon system performance. Similar success has been gained in modelling transport systems and in relation to water resource projects. However, the difficulty of modelling increases when the output depends more significantly on human behaviour; as in the library context. This difficulty has also proved frustrating in those marketing problems where performance depends on consumer response. Work in the hospital field where I have done most of my work provides interesting examples of situations which system performance is concerned with physical and economic phenomena on the one hand and with psychological factors on the other. At the physical end of this spectrum are the innumerable models of out-patient waiting time produced since 1950. Economic models have also been developed. In these, patient throughput, expressed in terms of standardised treated cases, is considered as a function of resource inputs such as medical and nursing staff levels. At the physiological level Howland (1960) suggests an adaptive control model in which a hospital is regarded as a system for controlling and regulating parameters like body temperature, heart rate and blood pressure. Finally Revans (1961) develops a model based on psychological theory whereby performance, considered in terms of morale, is related to the degree and nature of communication between people in the hospital. He, in turn, relates morale to certain indices of the effectiveness of treatment of patients such as length of stay and crude mortality rates. The relative abundance of physical models as compared with those concerned with patient health and welfare suggests that the difficulties of modelling these latter factors have forced work to concentrate on perhaps less relevant aspects.

The final group of models is concerned with the way in which a plan and the environment interact and in particular with the nature of the public response to the plan. Models of this type tend to be little more than assumptions. For example, in studies examining flood control schemes it is usually supposed that farmers will use all land freed from the danger of flooding in some optimal way. Again in road projects very simple and apparently arbitrary assumptions are usually made about the amount of traffic which will be diverted from existing slower roads. Rarely is any market research undertaken to gauge the likely community response to social plans. This is surprising when considered in relation to the cost of many social plans and when contrasted with the attention given to market research in the commercial field. The response of people to a plan depends largely on the existence of unsatisfied needs. These needs are translated by complex processes into specific demands for service. In particular the demand is likely to be a function of the type of service provided. For example, in relation to the demand for medical care, a specific request for service depends on an individual's assessment of his state of health and his knowledge of the type of services available. Attempts to provide services to meet popular expectations are very likely to stimulate almost insatiable demands as the iceberg of unsatisfied need is revealed. This is demonstrated particularly clearly in the health field by extensive studies carried out by Logan into the need for medical care.

III. MEASURES OF EFFECTIVENESS

Once the stage has been reached at which we have models predicting some of the relevant outputs of rival plans we require to determine which plan is the best. Since it is unlikely that any one plan will be best in all respects, an attempt is usually made to develop a criterion of overall effectiveness. The need for such a measure has, however, been questioned on the grounds that, since the decision maker is the only valid source of the values to be associated with each output, it would seem less devious to allow him to choose directly between the plans. Whilst this may be true in situations involving a strictly limited number of alternative plans and outputs it will ordinarily be difficult to present the relevant information in a form comprehensible to the decision maker when the number of plans and outputs is large. Thus in these cases at least a measure of effectiveness is probably essential and even in the simpler cases the decision maker may well prefer to have the output information summarised as a single number.

Accepting, therefore, the desirability of a measure of effectiveness I must consider the difficulties and feasibility of constructing one. As mentioned in the previous section some of the outputs of social systems have proved difficult to quantify and quantification would seem a necessary preliminary requirement to the development of a measure. Further it may well prove difficult to determine the true values held by the decision maker in relation to the various outputs especially if he must provide precise numerical statements himself. A good example is reported by Ackoff (1962) when he describes a problem concerning the planning of a dental service. He takes the objective to be to maximize dental health for the maximum number of people at a given cost. Unfortunately no-one including dentists, has managed to define dental health and therefore it proved impossible to specify a single output as its measure. The inputs and outputs which thus had to be used to characterise the service were,

(a) The various types of dental ailment to be treated.
(b) The number of people treated for each type of ailment.
(c) The cost of the service to those who are treated.

The problem was to determine what services should be offered and what charges should be made. In order to do this it is necessary to determine, for example, whether it is better to clean the teeth of one thousand patients or to repair the cavities in three hundred and clean teeth in five hundred, assuming equivalent costs. Likewise, the relative value of providing dentures to senior citizens or offering orthodontia to junior citizens has to be assessed. Although Ackoff states that he managed to devise ingenious methods for comparing plans he indicates that it proved impossible to combine the numerous outputs into a single measure.

At Lancaster a similar problem has been tackled concerned with the deployment of police patrols. The problem, in general terms, considers the positional distribution of patrols and the number of patrols of various types to deploy at different times of day. The essence of the problem is that the possible outputs of patrol activities are numerous and varied: from preventing murder to detecting breaches of speed regulations. Therefore, the differing outputs needed to be weighted.

In order to tackle the problem several surveys have been undertaken to obtain (from public and police) subjective valuations of a wide range of police achievments (Chambers 1968). It should be noted that the feasibility of obtaining meaningful values for some types of factors, by any means, has been questioned by some writers. Thus Hitch (1953) in considering how best a family should spend its income doubts whether it is possible to obtain the utility associated by the family with various goods and services. He states, "we could not write down the family's general utility function because the family could not tell us what it was, and we could not conceivably derive it from any other source". Two methods which have been suggested for assessing values are firstly to derive them by observing market behaviour and secondly, or alternatively, to estimate them by simulating the decision environment.

Since, in the social field, the market-price mechanism, if it operates at all, does not operate freely it is impossible to obtain market values directly. If, therefore, it is the market values that we seek they must be imputed in some way. In some cases this may be relatively simple; thus, for example, the water provided to irrigate previously unproductive land may be taken as the market value of the crop produced less the costs incurred in its production. In other cases, it is much more difficult to develop a reasonable argument to support imputed market values. A case in point concerns the loss of recreational fishing waters which has been variously assessed as equal to the market value of the fish caught or the wages the anglers could have earned by working instead of fishing.

Several specific problems of evaluation have received a great deal of attention in the literature and by way of illustration I will consider the value to be placed on time saved by improvements in transport systems. All evaluation procedures suggested, make the doubtful assumption that one minute saved on each of sixty occasions has the same total worth as a single saving of sixty minutes. They also usually distinguish between working time saved and leisure time saved, valuing the former at the recipient's normal wage rate and latter by noting that leisure time can be substituted for one of the following:

(a) Wages earned by workers.

(b) Transport expenditure, by those who travel in their own time and are able to choose between faster, more expensive, modes of travel and slower, cheaper ones. (see Moses and Williamson 1963)

(c) Housing and transport expenditure, by people who can choose between a more expensive house near to their work and less expensive house further away, (see Mahring 1961).

I feel that the deficiencies of each of these methods will be self-evident. The fact that time saving is often assessed at up to fifty per cent of the total benefits accruing from a transport project, however, emphasises the desirability of a more accurate method of evaluation.

Apart from the difficulties of finding satisfactory methods for imputing market values there are a whole range of other practical drawbacks encountered in estimating them. For example, the market values of some factors may be expected to rise over the project life, even relative to the general level of prices, but it will inevitably be difficult to predict the precise nature of the increase. Again a plan may be large enough to itself affect market prices but the extent of this influence will usually be difficult to gauge. Also the market behaviour observed in the private sector will rarely be fully competitive so that an imputed value may not be based on true market values. A particular case of a divergence between market prices and social costs and benefits derives from taxes on expenditure. Correcting for these shortcomings, whilst necessary, obviously allows scope for subjective bias.

Despite the enormity of the difficulties in the way of imputing market values we feel that, a more fundamental question that has to be asked, is whether market values are likely to adequately represent a particular decision maker's viewpoint. Whilst they may do so, it seems unlikely to take an extreme case, that a decrease in prostitution resulting from a social plan would be costed against the plan by many decision makers, let alone at its market price.

The alternative to utilising market values is to obtain values by studying choices made by the decision making group. Although the principle of this approach is applicable to the study of real life decisions the specific techniques developed are more appropriate to the analysis of verbally expressed preferences. The approach is typified by the method of Churchman and Ackoff (1954) in which subjects first rank the alternative factors and then place provisional relative values on them. These initial values are subsequently tested by presenting the subject with choices between groups of the factors and the values are refined if they are at variance with the choices made. The critical difficulty of all such approaches is the uncertain reliance that can be placed on verbal statements of preference or choices made in unrealistic situations. Nevertheless, this approach was used in the police project mentioned above.

IV. COMMENTS

In applying Operational Research to problems of social planning several difficulties have been identified: those concerned with modelling, especially when the performance of a plan depends significantly on human behaviour, and those concerned with measuring effectiveness, which appears to be especially difficult in social systems because of the non-monetary nature of some of the key outputs.

I consider that progress in modelling will only come from detailed studies of particular facets of total planning problems. I feel, however, that whilst this work is being undertaken, great care must be taken to ensure against undue emphasis being given to objective predictions derived from partial models as compared with the subjective ones obtained from relevant experts. That there is a real danger here will be evident to anyone familiar with the delightful oversimplifications often evident in the papers used in making high level decisions in Government Ministries.

Whilst therefore accepting the need for much low level work to be undertaken in order to develop models, I am not at all convinced that this approach avoids the real difficulties of measuring effectiveness in several of the community services. This essential difficulty seems to me to stem from the problem of properly measuring outputs like level of health, the extent of law and order and the standard of education, or of library service. These problems are present whe whether we consider, for instance, for health service planning, the health service as a whole, a particular hospital or an individual general practice. If Operational Research is to make a valid contribution to the central issues involved in social planning these problems of measuring non-monetary outputs must be solved.

V. REFERENCES

ACKOFF, R.L. (1966). Adaptive National Planning. Paper to Conference "Science in Planning", Department of Operational Research, University of Lancaster.

BEER, S. (1962). An Operational Research Project in Technical Education. Operational Res. Quart. Vol. 13, No. 2. 179-199.

CHAMBERS, M.L. (1968). "Patrol Deployment and Effectiveness". Department of Operational Research, University of Lancaster.

HITCH, C. (1953). An appreciation of systems analysis. J. Operations Res. Soc. Amer. Vol. 3, No. 4. 466-481.

HOWLAND, D. (1960). The development of a methodology for the evaluation of patient care. Engineering Experiment Station. Ohio State University.

LOGAN, R.F.L. (1964). Studies in the Spectrum of Medical Care. In "Problems and Progress in Medical Care". The Nuffield Provincial Hospital Trust, Oxford University Press. 3-54.

MAHRING, H. (1961). Land values and the measurement of highway benefits. J. of Polit. Econ. Vol. LXIX.

MOSES, L.N. and WILLIAMSON, H.F. (1963). Value of time, choice of mode, and the subsidy issue in urban transportation. J. of Polit. Econ. Vol. LXXI

REVANS, R.W. (1961). Standards for Morale. Cause and effect in hospitals. The Nuffield Provincial Hospital Trust.

SARGEAUNT, H.A. (1966). Planning for the allocation of resources in the Police Force. Paper to Conference "Science in Planning". Department of Operational Research, University of Lancaster.

WEST-CHURCHMAN, C. and ACKOFF, R.L. (1954). An Approximate Measure of Value. J. Oper. Res. Soc. Amer. Vol. 12. No. 2., p. 172-187.

Discussion

Mr. Dammers said that he found Mr. Ford's paper very interesting because it reflected his experience of actually doing the job of data collection. He added that data collection was very important and, referring to Mr. Duchesne's paper, he said that failures in data collection had held up the development of management information systems. Mr. Dammers said that Mr. Duchesne expected that in libraries the computer would be a clerical tool and if he had meant to imply that computers could not do more than clerical operations then he disagreed with him. Mr. Duchesne replied that he had not meant to imply that, but computers had to start somewhere and at first they would perform routine clerical operations. Dr. Urquhart said that libraries had become entranced by computers mainly because they did not know how to run their affairs on quite simple management techniques. He said that NLL had a management problem: it was necessary to ensure that the photo-copies went off promptly. The answer was quite simple: it involved someone going round every evening making sure that there was nothing awaiting photocopying. This was efficient, and much cheaper than any system involving a computer. Librarians must make sure that automation was used in the right way, and they must collect the right sort of data; it did not really matter if a library had 500, 000 or 750, 000 volumes, the important statistics were those relating to the needs of the users of the library.

Mr. Fairthorne said that Mr. Ford's paper showed that when carrying out a survey it was important 'not to disturb the animal', and if a survey was to be carried out in depth, then usually a pilot survey was necessary.

Mr. Dammers said that it was important to look into the future and visualize the potential uses of computers. Mr. Duchesne said that he had not meant to imply that it was not necessary to look into the future, or to consider the implications of things like on-line access in libraries, but it was important to remember the time scale of when this would happen. There was a big gap between what was technically possible and what the average library could afford: in Birmingham in 1972 the University would have the largest British-made machine at present on the drawing board, but the maximum number of consoles on line to it would only be 30; this was very few in the context of a large university. He agreed that it was necessary to plan for the next 5, 10, 20 years, but at the same time it was necessary to be realistic about what libraries could afford. Simple models were essential so that the librarian could understand the system: it was no good designing a system to collect masses of data which took up computer space but which the librarian would never use.

Referring to Mr. Ford's comments on interviewing, Mr. Vickery said that although he acknowledged the shortcomings of the interviewing technique as a method of collecting data, interviews could often reveal facts which would be missed by instant diaries and questionnaires. Librarians as managers needed two sorts of data: survey data for particular investigations, and routine control data; an efficient management information system should be able to give both types. The control data was needed so that the librarian could see how the library was functioning, e.g. whether there was a backlog of books awaiting cataloguing. Mr. Dammers agreed with this and said that unfortunately sometimes computers had been introduced to deal with muddled situations where the objectives had not been precisely defined; in such cases the computer only made the situation worse. Dr. Urquhart said that the management problems in librarianship were basically so simple that usually the introduction of a computer was not necessary. Mr. Snape observed that Dr. Urquhart had a closed access library which was simpler to run than an open access one. Dr. Thompson agreed: it was much easier to obtain usage data when a slip was filled in for each issue than in an open access library which had almost no control of, or even data on, in-library or reference use of books.

Mr. Buckland said that there was a Pavlovian response amongst librarians, who seemed to assume that research on library's problems necessarily implied questionnaires: he was proud of the fact that during the thirty months of the research project at Lancaster only 14 questions had been asked of the readers. It was important to decide first which data were relevant to a particular management decision and only then to decide how to collect the data: once you knew the precise information you wanted it was often quite a trivial matter to collect it. He described an experiment which was carried out at Lancaster to collect information on the use of scientific periodicals by use of booby-traps.

Mr. Brookes said that one aspect of data collection which interested him was that often he had to analyse data from a computer print-out; this took a long time and often the analysis could have been completed in a few seconds by the computer if it had been incorporated in the programme. One example of this was the fitting of Bradford-Zipf curves to MEDLARS data: he regretted that there was no programme for this incorporated in the MEDLARS system. Dr. Urquhart said that the purpose of MEDLARS was to provide references in reply to questions, and not to turn out Bradford-Zipf curves; there was, however, a programme at Newcastle which listed references under periodical titles. Mr. Brookes agreed that the purpose of MEDLARS was not to turn out Bradford-Zipf curves but said that a few years ago, when the organizers of MEDLARS wanted information about the productivity of the periodicals scanned, they had to obtain this information by clerical means, and he found it surprising that the system did not provide such control information automatically.

Dr. Thompson said that librarians often thought that the only way they could analyse data obtained from surveys was by a computer; but the data which librarians gathered could usually be sorted mechanically and this was much cheaper than computer sorting. Dr. Urquhart said that at the NLL mechanical sorting was used when the number of cards involved was of the order of a few thousand, and computer sorting was only used when the number reached the hundred thousand mark.

Mr. Duchesne said that there was a tendency for librarians to think that if they recruited a systems analyst then they could delegate all their problems to him. Often the systems analyst made a good job of designing the computer system but if the librarian has not told him the right questions to put to the computer then the system would be useless.

Miss O'Connor said that often when the results of library surveys were published not enough background information was given to enable the reader to extrapolate the results to his own library; one example of this was giving the issue figures but not the number of registered undergraduates.

Dr. Thompson asked Mr. Mackenzie what considerations he had taken into account when he had decided on a central university library at Lancaster, as opposed to a number of departmental libraries. Mr. Mackenzie said that he had not used O.R. techniques, but had looked at older universities, with fragmented departmental libraries, and this seemed to be a very expensive way of providing a library service. He said that on the whole there had been no moves from departments to set up their own collections, so they were probably satisfied with the service; even when there were 8, 000 students at Lancaster the farthest point from the library would be less than five minutes walk under cover. One of the advantages of the college system was that there was a good deal of traffic past the library during the normal daily activities of students and staff, which made it very convenient for them to visit casually.

Mr. Mackenzie asked how was it possible to evaluate the benefit gained from the various library services; e.g. the value of one inter-library loan as opposed to ten books issued in the library. Dr. Hindle agreed that these were fundamental questions which were difficult to answer but tham some extension of the technique which he had just outlined might be applicable. Dr. Thompson said that it was necessary to know how important expenditure was in relation to the success in achievement of various objectives. Mr. Morley said that in defining objectives it was important to get the opinion of the user. Mr. Mackenzie agreed with this but pointed out that the librarian didn't always want improvement in the same things as the user; it was necessary to define objectives first; he cited as an example of this the Chicago library staff survey of automation objectives. Mr. Morley said that one way of analysing the expenditure was to measure the output from certain activities. Mr. Dammers said that one factor was that of the user's time: the cost of delay to the user must be incorporated with the cost of borrowing an item from outside. Dr. Hindle agreed that this was often neglected and that more work was needed here. Mr. Dammers said that it should be possible to attribute a definite value to the delay in getting material for a research worker; the speed of access to literature had implications for research programme development - the more difficult the access the greater delay in the programme; delay would also break the user's train of thought. Mr. Mackenzie said that the delay (at least in universities) was often filled in with other activities which the research worker had to do and that this penalty might in some circumstances be illusory. Mr. Brookes disagreed: when scientists did not get the specialist information services which they were used to in industry, the nature of their research was often amended, becoming more fundamental and futile.

SESSION 6: REVIEW

B.C. Vickery	Library management research; what next?
F.F. Leimkuhler	Planning library services: an overview
Discussion	

Planning library services (University of Lancaster seminar)

Aspects of library management research

B.C. Vickery, Aslib

The title of this seminar clearly indicates the purpose of library management research - to aid the planning of library services. If we look at the matter in its simplest terms, there are four services that a user can expect from a library:

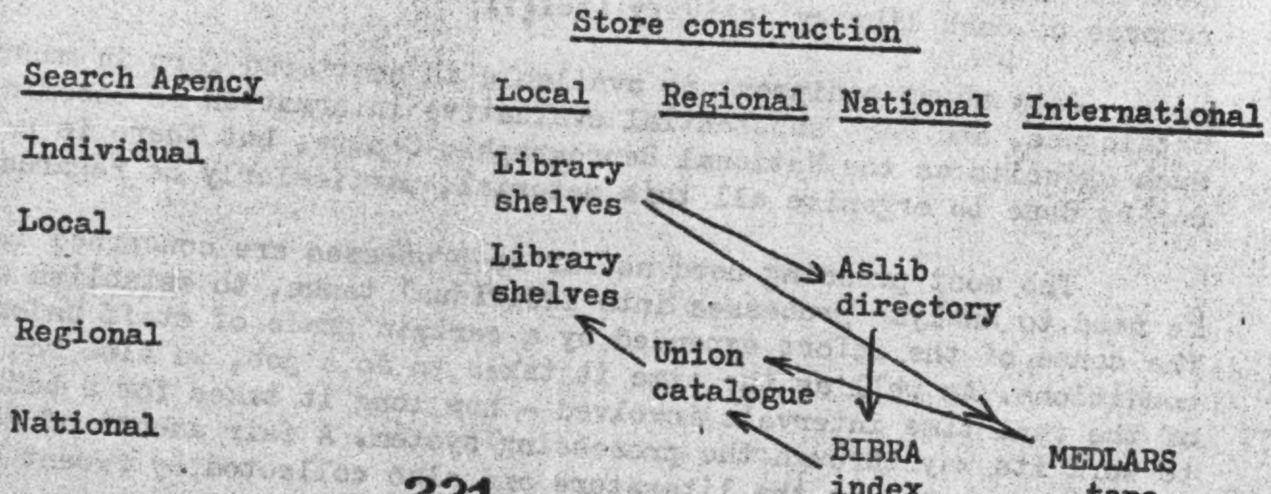
to provide documents from stock
to provide references from a file
to provide information from documents
to provide locations of documents, stock or files.

Each of these services involves both storage and retrieval. On the one hand, stocks, files, documents and location lists must be constructed - this is information storage. On the other hand, they must be searched - this is information retrieval. These two phases need not be, and usually are not carried out by the same people. So we have eight basic activities:

	Storage	Retrieval
Document	x	x
Reference	x	x
"Data"	x	x
Location	x	x

These activities can be carried out by various "centres of action". At one extreme, all the retrieval activities can be performed by the individual library customer. He does not perform storage activities(except in his potential capacity as an author). The construction of document, reference, data and location stores may all be the function of institutions. These institutions can be local, regional, national or international, and can vary in subject coverage.

A simple illustration will help to make my point. An academic researcher wants some information on the toxicity of certain chemicals. He searches the (locally constructed) library stock. He interrogates his local librarian, who scans the (nationally constructed) Aslib directory, and suggests an approach to the British Industrial Biological Research Association and MEDLARS. The Association searches its index, UK MEDLARS searches its (internationally constructed) tapes. References from these searches are sought in the Regional union catalogue, and documents provided by other local libraries. The procedure can be outlined as follows:



The purpose of this simplified analysis is to suggest the basic problems that face library management.

(1) Looked at from the national point of view, there is the problem of the allocation of activities among centres. With regard to each of the eight storage or retrieval activities, what is the best distribution of resources among local, regional, national and international centres? What degrees of subject

specialisation are appropriate in each case? Here we are considering the optimisation of the hierarchical library system.

(2) From the viewpoint of an individual library, there is a similar problem of resource allocation. Given the existing library system, what stock will be locally provided? What reference files and location indexes will be locally constructed? To what extent will external stores be used? How much searching will be left to the individual customer?

(3) Having decided on the activities to be undertaken locally, each library has the problem of organising each of them optimally.

(4) Once in operation, library management needs to monitor activities continually so as to be able to react to change.

The solution of these problems is the task of systems analysis, design, and control: it is the development aspect of library management. Beyond this is the research aspect. It seems to me that library management research has three tasks: (a) to provide background data about system components; (b) to establish relations between system variables; and (c) to develop new techniques of systems analysis.

Data about system components

There are several groups of components in a library system: documents, of course, and records of documents; equipment and accommodation; the processes that weave these together; operators (library staff); system aids (manuals, rules, etc.); and users. Factual data - both qualitative and quantitative - about the characteristics of possible system components is needed to aid the systems designer to choose among them.

Librarians already know a good deal about the characteristics of documents. The physical bibliography of the bound book is being extended to cover physical characteristics of other media such as microfilm. Information about the varied characteristics of origin (author, publisher, etc.) is enshrined in our cataloguing rules. The characteristics of subject literatures (volume, rate of growth, forms of publication, links with other subjects, etc.) have been explored haphazardly and need continual updating as the field of knowledge develops.

Although library activity makes immense use of every kind of bibliographic record, only the catalogue entry has had much systematic study. Recently - stimulated by the MARC project - research into records has increased. At Aslib we have made an analysis of the bibliographic and administrative data elements used in records of each library activity (ref 1).

Data about equipment is available in scattered form in manufacturers' catalogues, and more substantial evaluative information is becoming available from such agencies as the National Reprographic Centre, but there is no doubt much that can be done to organise all this material, particularly as regards relative costs.

The most pressing need as far as processes are concerned is once again costs We need to analyse processes into individual tasks, to establish unit costs - in the sense of the effort expended by a certain grade of staff under certain conditions. As well as the time it takes to do a job, we also require knowledge of the real time intervals involved - how long it takes for a book or a request to work its way through the processing system. A fair amount of such information has been reported in the literature and also collected by recent UK research projects.

Lastly I will mention users. Studies of user behaviour, stated wants and needs have proliferated in recent years. Despite much criticism of the quality of these studies, a good deal of knowledge that was previously impressionistic has been quantified. We are now moving into a period of specialist studies - more specific user groups, particular kinds of information or documents. Of particular value is the study of user failures: why they do not find what they seek is the cost to the user of using the library - recognising that we delegate most retrieval activities to him.

Relations between system variables

Some relations between variables in a library system are no more than roughly quantified common knowledge - for example, scientists use libraries more than engineers. Other relations may take the form of complex mathematical models - as in Buckland and Woodburns' study of the implications of scattering and obsolescence.

Since a library system comprises many components, and many characteristics of each component may vary, and many of these variables may interact, the scope for research in this area is very wide. As an illustration of the complexity, I have in Figure 1 set down over two dozen variables that may affect the success of user search for documents in a library, and hypothesised interactions between them. For example, search success may be affected by the modes of approach the user adopts (author, subject, title,etc.) What modes are used may depend on what modes are provided on the shelves or in secondary access tools (catalogues, bibliographies, etc.). The modes provided may depend upon the library's estimate of its users' needs, and this in turn on the methods it adopts to estimate needs.

How is research to choose what relations to investigate? Clearly, three criteria can be adopted. First, it must be possible to categorise (preferably in a scalar or quantitative way) both the influencing and the dependent variable. Second, the influencing variable should be one that library management can in fact manipulate. Third, the dependent variable should as directly as possible relate to the services that the library exists to provide. To know that scientists read more than engineers is useful background information for the librarian, but he must accept the users his environment offers. To know that discarding all literature more than twenty years old will only affect his users' success rate by one percent (or whatever the figure might be) can directly influence discard policy.

There is a danger, however. As soon as we have quantified a particular relation, established a particular model, there is a natural tendency to optimise on it - e.g., to shelve books most economically by dividing them into 2 or 3 height sequences. But because of the complexity of interactions, suggested in Figure 1, this may not be optimal for other library objectives - to have several book sequences may lower search success.

Techniques of systems analysis

In Figure 2 are listed a couple of dozen techniques available to the library systems analyst. Most of them have been developed in the general management field, though a few are native to libraries (e.g. in section 1, citation analysis; in section 5, statistical bibliography). Even to apply standard analytical techniques to libraries is a form of developmental research, for they need to be adapted and tested. At Aslib we have been evaluating a technique of data collection first tried out at Lehigh University (ref.2).

Library management research needs to be particularly active in developing measures of effectiveness. The importance of even a simple measure of user success has been stressed at this seminar. Several types of measure have been employed: (a) What proportion of potential users do use documents, (b) what proportion of sought documents are found, (c) is needed information obtained from documents supplied (relevance assessment), (d) is the user a better man for having the information (assessing user performance or preference).

Evaluating the effectiveness of our services is indeed the crucially difficult problem facing library management. It becomes more difficult as we widen the objectives of the library. Consider the following hierarchy of objectives: (1) to acquire one copy of every document published on, say, toxicology;

(2) to meet without delay 90% of requests for documents on toxicology; (3) to satisfy 90% of user needs for information on toxicology; (4) to provide optimum information backup to those working on toxicology. There is no guarantee that meeting an earlier objective in the hierarchy is either necessary or sufficient to satisfy its successor. Each successive objective is nearer to what we think of as the true purpose of a library, yet it is successively more difficult both to specify the objective in detail and to determine whether it is attained. It is in this area of evaluation that new research ideas are most urgently needed.

In conclusion

I have presented a view of library management research as a back-up to the practical tasks of library systems analysis, design and control. To my mind, it includes descriptive analysis to provide data about system components; factor assessment to establish relations between system variables; and the development of analytical and evaluative techniques. I have suggested areas where research is needed. The topic is now open for discussion - and action.

References

(1) P.A. Thomas and H.East "The use of bibliographic records in libraries", Aslib 1969 (Part 2 in draft).

(2) "Analysis of library processes" Aslib Research Dept., unpublished.

	Dependent variables	Influencing variables																										
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	Search success		x		x	x	x	x	x	x	x	x	x				x	x										
2	Personal characteristics																											
3	Work situation																									x		
4	Search time available		x	x																								
5	Knowledge of store organisation		x	x				x																				
6	Modes of approach		x	x	x	x			x		x						x											
7	Advisory staff available																								x			
8	Sequences in collection														x	x	x	x										
9	Display guides to collection								x			x		x			x											
10	Modes of secondary access								x					x	x		x	x										
11	Physical housing of collection													x		x		x	x									
12	Physical form of access aids										x			x														
13	Estimate of required access speed																							x				
14	Estimate of modes of approach																							x				
15	Restrictions on store arrangement																	x	x									
16	Size of collection																		x	x	x	X					x	x
17	Scope of collection																			x	x	x					x	x
18	Restrictions on store size																				x					x		
19	Availability of acquisition aids																				x						x	
20	Skill of acquisition staff																								x			
21	Acquisition policy																						x			x		x
22	Estimate of literature demand																							x				
23	Methods of estimating needs																									x		
24	Staff allocation policy																							x		x		
25	Type of environing organisation																											
26	Dissemination activity of publisher																											
27	Range of published literature																											

Notes

Notes	
1	as variously measured
2	e.g. age, education, psychological type
3	e.g. job, rank, project stage
6	e.g. by author, title, subject
10	e.g. catalogues, indexes
11	e.g. shelves, cabinets
12	e.g. cards, punchcard, magnetic
15	due to layout etc.
17	scope as to subject, bibliographic type, literary type, physical form
19	e.g. publishers' announcements
21	including discard policy
25	e.g. industrial, academic
27	as 17

Figure 2

Some analytical techniques for investigating libraries

1. Users, customers

   - Observation of behaviour
     - (a) direct
     - (b) indirect, via records of behaviour
       - (i) self-compiled (citations, diaries, questionnaires)
       - (ii) compiled during system operation

   - Collection of opinions, statements of wants
   - Analysis of needs
   - Interaction with system
     - Measures of effectiveness (all kinds)
     - Queueing theory

2. Processes
   - Work and method study (including work simplification and procedure study)
   - Value analysis
   - Network analysis

3. Operators
   - Work measurement
   - Activity sampling
   - Motion study
   - Performance analysis
   - Job study
   - Organisation theory

4. Equipment and accommodation
   - Replacement analysis
   - Layout analysis
   - Ergonomic analysis

5. Documents, records, collections, files
   - Statistical bibliography
   - Forms design
   - Store and file organisation
   - Code design

Store and file organisation

6. General
   - Statistical methods
   - Sampling
   - Charting and graphical methods
   - Quality control
   - Resource allocation
   - Simulation
   - Standardisation
   - Data processing

Planning University Library Services: an overview

By: F. F. Leimkuhler,
Purdue University,
U. S. A.

University libraries are growing at a steady rate according to a rigid manual technology that projects proportional increases in all inputs without a clear understanding of the economic advantages and indirect costs of continued development along these lines. The commitment to collect everything that is relevant to a very heterogeneous group of users and to hold large collections indefinitely; the need to be consistent with past procedures and national standards; the absence of good marketing mechanisms, the lack of funds to experiment with new methods; and the example of larger libraries elsewhere all seem to indicate that university libraries will continue to grow along fairly conventional lines. Because this projection calls for exponential increases in both material and personnel costs, libraries can expect to be even harder pressed in the future to devote all available resources to the maintenance of programs which offer little promise of fundamental innovation. Indeed the larger the scale of operations the more difficult it is to foster innovations.

Some relief is anticipated through the establishment of national, regional, and commercial services which can provide some additional economies of scale. But there is no firm schedule for these services nor hope that they will do much more than replace existing in-house activities. The absence of or delay in substantial library innovation is encouraging the proliferation of small scale information services which compete with the libraries for funds and user support. These services are in a better position to exploit newer technologies and develop specialized user markets. Such competition, while stimulating innovation, can lead to serious inequities in the provision of information services to the academic community and actually distort the collective mission of the university.

There is a need to find new ways in which the traditional mission of the university library can be sustained and yet allow for a much greater degree of flexibility and resourcefulness in developing and exploiting technical innovations. This would allow the libraries to respond positively to the conforming pressures of interlibrary networks and the changing demands and evaluations being made by library users. The libraries need to be able to bargain effectively with the new centralized services being made available to them, and this bargaining power must be based on the strength of their own technical competence and professional responsiveness to the university community. At the same time, the libraries must be able to compete with the new kinds of specialized information services being offered to the users on an ad-hoc basis.

University libraries should be able to draw considerable economic advantage from their size, reputation, and monopolistic position on the campus. Furthermore, their proximity to university research programs should provide them with a privileged access to the latest developments in information technology. To take advantage of these opportunities, the libraries must develop the organizational competence that can enter into experimental ventures and carry out effective market research to justify the introduction of new services. It is important that libraries re-examine their planning doctrines so as to justify and secure the additional funding these ventures will require both by augmented budgets and by diverting existing funds into more innovative programs. The recruiting of research personnel and the indoctrination of present personnel to this end is no small part of the effort needed for a farsighted program of library development.

The Computer and the Book

The technical issues in library development revolve around the computer and the book, two of man's more remarkable inventions. In looking back over the past one hundred years of library development it seems clear that the quest has always been for greater comprehensiveness and accessability of collections. The ultimate goal seems to be instant access to all of the world's recorded knowledge. No library is so big that it does not anticipate doubling its size in the next ten to twenty years. No library is satisfied with its ability to satisfy all user requests quickly and completely. It should be no surprise that some prophetic persons see in the computer the realization of this goal and anticipate the day when a giant electronic brain will provide man with instant access to all that is worth knowing. But that day still seems so far away that it does not occupy a significant place in plans for libraries of the immediate upcoming generation.

The computer does create an immediate challenge to libraries, not in its ability to replace the book, but in its ability to replace the catalog. Before the computer, the main technical issue in libraries was between the book and the card catalog, that is, between the book and records about books. Books are handy devices when used individually, but collectively they are cumbersome, difficult to manipulate, and wasteful of space. Manual card records are easier to handle in large numbers, and one can afford to file similar records under different headings to allow a variety of access points to the same store of information. Like books, cards make manual random access to any point in the file relatively easy for both purposes of retrieval and for the inter-filing of new items. However, this ease of access and ability to accomodate steady expansion is obtained at the price of a rigid conformity to common rules and lack of adaptation to new needs and circumstances. The economies achieved through standardized procedures are eventually offset by the increasing cost of lost opportunities to provide specialized services as the need arises.

Computer-based information systems offer the advantage of ease and economy in the manipulation of machine-readable records. For large record systems like libraries it costs more both to enter new information and to store old information in machine-retrievable form. Technical advances in both data origination equipment and storage equipment should reduce these costs substantially in the long run. Operational changes which minimize the amount of original entry work and make possible shared use of large data banks will also reduce the costs substantially. The net effect of these reductions should be to force the cost per record down to a point which is competitive with manual record systems. Even so, the library cannot expect to save money on a cost per record basis by adopting computer methods; the advantages must come through the computer's ability to process data at fantastic speeds and to manipulate records in a manner which would be unthinkable by manual methods. How to translate these opportunities into enough tangible benefits for library patrons and long run economies in library budgeting so as to justify the introduction of computer technology, is a difficult task: it took 50 years to win acceptance of thetypewriter as the standard way of producing catalog records.

There are perhaps two positive approaches to the introduction of computers in the library in addition to the obvious negative one. A conservative approach would be to encourage the development of systems which emulate existing manual systems and compete favourable on a cost basis. This is the "cardless catalog" approach and it offers perhaps the least mental resistance to change on the part of the library profession. However, it runs the risk of slavish adherence to certain features of manual systems which are unnecessary in a mechanical system. A more optimistic approach is to begin with a thorough re-examination of the purpose and functions of libraries. This is the management systems approach, that is, by working backwards from the ends to the means of accomplishing given objectives subject to certain criteria of performance and limitations on inputs. Where the necessary technical detail is unavailable for completing some juncture between an output and an input, a "black box" is substituted and earmarked for future study and development. This permits one to continue to sketch in the larger plan of a hypothetical system and identify those key points where basic research is called for. In practice we can expect to see developments on both fronts. Both approaches are valid and should be exercised fully at all levels of the development effort with one as a check on the other. Libraries should seize every opportunity to make use of the computer in order to familiarize themselves with its use. It is likely that university libraries and computing centres will be merged someday under the same management.

The Problems of Size and Retrieval

Operations research concentrates on problems of innovation and transition. The use of a team of analysts with different scientific backgrounds and operational experience is intended to provide a balanced, objective view of problematic situations. The focus is on problems and the main effort is to define a set of research problems which closely approximate the real world problems of a particular decision-maker. This process of abstraction permits the use of analytic methods and insights from the entire world of applied science. Problems in production scheduling may suddenly have important ramifications for problems in book processing, queueing models may provide important insights for loan and duplication policies, military search theory may apply in certain ways to the problems of information retrieval. Model building in operations research begins and ends with problem formulation. A model is a problem statement in solution form and is thus the most precise way of stating a problem. The search for appropriate models which capture the essential features of problematic situations is the main thrust of operations research.

Some attempts have been made to define "the library problem" as the problem of information storage and retrieval or more specifically as the problem of document storage and retrieval. With document storage there is a problem of size and space. Partly it is a problem of growth, and partly it is the problem of arranging items in space and time in a relatively efficient manner. Decimal classification is a model for solving the book shelving problem. It is an orderly, linear way of using space which works reasonably well in open-stack storage, but is not necessarily the most efficient way of using space in all circumstances. It has its drawbacks as does any idealized model. Its chief advantage is that it allows for the indefinite interspersion of new acquisitions while maintaining some kind of logical order. In practice it accumulates obsolete, redundant, and mismatched information which tends to inhibit search (but may give rise to serendipity). Orderly violations of the system are usually necessary for items with unusual usage patterns or unusual dimensions. Prime space is sacrificed and reshelving is required to allow for continued expansion. Discard policies are difficult since there is no age control, and the control of losses and mis-shelvings is also made more difficult.

All of these indictments of the decimal system are not new and have been considered before. While some modifications of uniform shelving have been made in certain situations, it still remains the standard for library practice. However, as libraries get larger and funds get tighter, the advantages of adopting unconventional storage patterns should be reconsidered both with respect to library costs and service to patrons. In recent years, theoretical approaches to problems of storage have received more attention and should provide more powerful methods of identifying and evaluating alternative storage schemes for libraries. These methods may find their first applications in the spatial arrangements of microform storage systems, since, while miniaturization provides an enormous reduction in the scale of a file, it does nothing to change the problem of arrangement except to make it even more necessary to make a precise mathematical statement of the storage policies adopted.

The use of unconventional storage policies and microform files usually deny direct user access to the collection and greatly increase the problem of retrieval. In fact, virtually every proposal for solving the problems of library size (inter-library networks, depositories, clearinghouses and central lending services, information centres, and miniaturization), depend heavily on the notion of remote access, i.e. the separation of the storage and retrieval functions. It follows that the highest priority must be given to the intellectual, technical, and social problems of providing remote bibliographic access to sources of information. The problem of retrieval is seen as the most important "black box" in the development of large-scale, automated information systems, and to some researchers it is considered "the library problem". The retrieval problem and related problems of awareness and selection have been a central concern of much of the basic research in library science for a long time. However, progress in this area of research has proven to be more difficult and more costly than originally anticipated and we seem to be many years away from large scale operational retrieval systems of the kind needed by a university library. Meanwhile, we can expect to see greater immediate use of the computer in exploiting the more limited opportunities available in the management of conventional library materials.

The Library and Its Users

In his recent book, The Systems Approach, C.W. Churchman argues that the user should be the most important consideration in the design of libraries and information systems. He is sceptical about the design of traditional libraries. He finds that they pursue programs without relating them to the needs and activities of the larger communities which they serve; and that libraries do not determine the true requirements and costs of their programs; nor do they evaluate programs in terms of the user's measure of performance, i.e. benefit minus cost. Churchman is also critical of efforts to measure performance in terms of physical activity and transactions rather than benefits. He says that the true benefit of an information system must be measured in terms of the meaning of information for the user, and the value of acquisition. Thus, the mere computerizing of a traditional library does not constitute a systems approach. For example, the computerizing of the Library of Congress catalog fails in this regard because the basic system questions associated with the catalog and its usage have not been answered by putting the catalog into machine-readable form.

To Churchman, the more important and difficult question in the design of information systems is what model of the user should be incorporated into the system. If behaviour is the guide to value, then the model must evaluate information in terms of user behaviour. But this raises serious questions about the implicit philosophic assumptions which must be built into such user models and into the systems which contain them. Even a speculative or "dialectical" information system which could examine its own premises and argue from different points of view, does not escape the spectre of dehumanization which Churchman finds implicit to the whole systems approach. The attempt to use human benefit as the formal criterion for systems design becomes largely self-defeating to the analyst, who deals in ideas rather than people. Churchman concludes his discussion with the remark that: "People are not 'better' just because they process information more rapidly or more coherently. People are better because they are better in a moral or aesthetic sense."

Because conventional libraries are essentially manual systems which depend heavily on the notion of direct user access to stored materials; because they deal primarily with books, which are large, coherent chunks of information bearing the personal stamp of authors; because most serious readers have been schooled in the use of these materials; and because libraries are staffed by trained professionals who make a personal commitment in vouching for the quality and integrity of their collections, there are many built-in controls which tend to keep libraries relevant to the needs of the users. There is always room for improvement and the possibility of outmoded practices, but the more important problem is how to expand the traditional system to accomodate the flood of new information and technology and still preserve a quality of relevance. It is in the design of new systems that can handle a many-fold increase in the flow of information, that we need to examine carefully the social context in which the system is to operate, especially with respect to fact retrieval systems, as Churchman points out. The promise of these new systems and the pressures for greater service make it necessary to try to use the systems approach to identify the critical problems which are likely to arise in the projected growth of conventional libraries.

One way to expand the concept of a library and imbed it in a system which includes the user as an active component is the model proposed by N. R. Baker. Baker depicts the library as part of a service system consisting of three principal components: the library itself, the community of potential and actual users, and an administrative or funding agency which supplies the library with resources in anticipation of user needs. By introducing the funder as the final arbiter of the proper use of library resources, Baker can examine library performance relative to user needs and university policies in a formal way. From a study of the interactions among the three components, Baker concludes that the library can expect to become increasingly constrained in its decision-making unless it can demonstrate its ability to make satisfactory decisions before users exercise their influence and before funders exercise their powers to force decisions in the directions they favour. Libraries must convince the others that learning, not influence and control, is the most productive approach to the solution of library problems; and that librarians are making full and effective use of the political, economic, and technical resources which have been made available to them. Furthermore, because of the behavioural conditioning which guides users in their evaluation of library services, the library must be quick to advertise any changes and improvements in the services offered if they wish to get actual and potential users to respond to these changes. The initiative must rest with the library.

Planning and Programming

By viewing the library as part of a larger system which includes users and funders, it can be seen that the problem of measuring the performance of libraries is first of all a problem of defining the context with which performance is to be measured. If the library is seen primarily as a servant of its users, then the users call the tune and costs become just a matter of getting as much money as possible to satisfydemands. The management of such a library would have to be particularly sensitive to the whims of those users who protested the loudest in order to allocate resources so as to minimize the general level of dissatisfaction that would prevail. If on the other hand, the library is seen primarily as the local branch office of an international system which has cornered the market in recorded knowledge, then local costs and local benefits would play a relatively insignificant role in the determination of operating policies. Local users would be offered service on a take it or leave it basis at a price which would insure the continuance of the system's monopolistic position.

Both of the above views are obvious distortions of reality, but are not entirely fictitious. A third approach is to view the library as an information service agency for the organization which funds it. Benefits and cost would be determined relative to the goals and resources of the larger institution which the library serves. The measurement of performance of a university library would depend on the performance measurements of the university itself, since the critical question in a decision to add one more dollar to the library budget would be whether or not it would benefit the university more than if the dollar were spent elsewhere. Exactly how the library spends the money or exactly which user will benefit directly from the expenditure is of secondary importance to the first test of benefit to the university.

In some respects, it seems that this latter approach merely exchanges one complex measurement problem for another, i.e., the measurement of university benefit instead of user benefit. It is conceivable even that university benefit might be defined as being the same as user benefit, in which case we would be back where we started. But there is an important difference in that the responsibility for validating the measurement of benefit rests with the university and not with the library nor with the users. This does not mean that the university must or should establish the appropriate measures of performance for a library. It would be better for the library to retain the freedom to do this, but the library should try to make its measures explicit and the university should approve those measures which are used in the justification of requests for funds.

This approach to library planning is most consistent with the current interest in using program budgeting techniques for university administration. Recently, in the journal College and Research Libraries, J.E. Keller of University of California points out that "sooner rather than later these techniques of program budgeting and cost benefit analysis will be applied to library operations." The idea behind the re-organization of budgets along the lines of specific programs is to facilitate the assessment of the benefits. In a program budget for the M.I.T. Libraries, R.Shisko divided library costs between two major programs: instruction and research. This makes it easier for university management to evaluate the budget relative to the goals of the university, in terms of both the absolute values for each program and the relative sizes of the various programs. The orientation is definitely in the direction of the university's goals rather than those of the library or its individual users.

While there is a tendency for program budgeting to make the library subservient to the university administration, it is not necessary for this to be the custom; and in order to avoid this the library must take the initiative in developing the manner in which library programs are presented and justified. There is no reason why the library should not shape as well as respond to the goal structure of the university. If library effort is reduced completely to supporting roles in teaching and research then the description of those missions should include a more precise definition of how the library is expected to contribute to them. If it is left to the arbitrary judgment of certain teachers or certain researchers and exercised after the costs are in, then the library cannot develop the kind of long run plans it needs.

Benefits and Costs

The search for acceptable ways of evaluating the benefits and costs of library services is complicated by the tendency of libraries to rest their case on a single total cost figure, with the implication that the internal problems of resource allocation are not subject to analytic scrutiny outside the library and only the total cost figure can make any sense relative to the goals of the larger university environment. The possibility of finding a single comparable benefit measure is extremely difficult if not impossible. Program budgets,on the other hand, tend to cut across established organizational lines and are even devisive in their effort to restructure the library into a large number of programs which are competing for funds and resources from a common limited pool. This makes it possible to focus attention on the development of cost-benefit techniques within the context of individual programs, and it is not uncommon to find most of the formal analytic work at this level. The larger problem of reconciling the differences between the measures of effectiveness for different programs is often left to the judgment of top level decision-makers, although the formal analysis should help in establishing program trade-offs.

At the program level the first concern is to establish the correspondence between valuable inputs and those outputs which contribute to the stated mission of the program. In general, one expects a greater degree of program success as more funds are devoted to a particular program. The important thing is to find out the marginal or incremental gains from added costs and to identify the different methods of accomplishing the program which are relatively efficient over some range of activity level for the program. Cost-benefit relationships which include shifts to new methods are needed for long-range planning, while those without such shifts describe the short-run possibilities. Short-run patterns tend to have relatively limited ranges over which they are most effective, and beyond which costs begin to rise steeply, relative to benefits, as program activity increases. The more difficult problems in planning are those situations where process changes are indicated and substantial amounts of new investment are required to effect the change and permit continued expansion of a program in an efficient manner. This is similar to the problem of whether or not to introduce new programs.

Library programs fall into two main groups: reader service programs and supporting technical processing activities. The service programs should have a strong voice in the evaluation of the support programs, since the former are more directly related to the goals of the university. In a similar way, the users of library services should have a strong voice in the evaluation of the service programs. From an analytic standpoint, user acceptance of a library service is reflected in its demand pattern, that is, in the relationship between the price users must pay for a service and the use which is made of it. "Price" is defined here as the effort required of the patron and it is directly related to the time required of the user and inversely related to the probability of satisfaction. Another important factor in the determination of a user's price is the value he places on time and success. For example, faculty and students probably apply different valuations to these factors; and by making the same service available to both without distinction, the library in effect prices the faculty member out of the market for a library service.

The library can reduce the user's price and increase the demand for service by increasing its own costs, and the important question in program planning is to know where to establish an optimal level of program activity. One approach is to find that level where the average cost to the library per unit of service is minimized, or the average number of service units rendered per dollar spent is maximized. This is a conservative approach which promotes technical efficiency in the expenditure of library funds, and under certain conditions it may achieve a near optimal solution relative to user expectations also. If it is coupled with a vigorous campaign to attract users to the service, this approach is not unlike business policies which promote the dual objective of producing at minimum average cost and selling all that is produced at a favourable price. However, serious inequities can arise from a policy based on benefit-cost ratios, such as the exclusion of certain users from needed services or the failure to provide necessary services to the university which may be relatively inefficient on a library cost basis. Hence, the need for continual refinement and auditing of program descriptions and the methods of analysis, as well as provision for intelligent intervention by higher decision-makers.

Some Final Remarks

Various problems in the planning of university libraries have been discussed in terms of what seems to be the most important issues at stake today. The main purpose in doing this was to try to identify those short-run approaches to the resolution of these problems which are the most promising for implementation now in the planning of libraries. However, such planning should be consistent with the most likely trends over the long run. With respect to certain problems a high degree of experimentation and adaptation is required. The main conclusion to be drawn from this review is that libraries must find a way to divert a significant portion of their budgets into such staff activities as planning, control, technical development, and market research. A ten percent budget allocation does not seem to be unusually high for these purposes, but the specific amount should be subjected to the same kind of program review as is indicated for service and processing activities. A final conclusion is that these library development programs should be staffed and organized so that they are in close communication with the entire university community.

Discussion

Mr. Morley asked if it was possible to put a value on the cost of a document: you either know the information in a document and therefore didn't want it, or you didn't know, and therefore couldn't find it; Mr. Mackenzie said that he had come to the same conclusion by a slightly different route: a person searching for information finds some which is relevant and some which is not, and it was only possible to deduce an average cost or value of all the bits of information he has collected.

Dr. Thompson said that when assessing the value of library services to readers the moment of truth came when it was decided to charge for a service; it was possible to cost on paper, for inter-library loans for example but this costing only becomes meaningful if the user is actually charged for the service. Nevertheless this practice would only determine the subjective value placed on such services, which might be quite different from their absolute value. Mr. Mackenzie said that about a year ago a coin-in-the-slot copying service was installed in the library at Lancaster and it was found that 90% of the users were undergraduates, who spent on average £100 a month on it. Thus undergraduates were prepared to pay for photocopying, although they could borrow most of the contents of the library.

Dr. Thompson said that the National Libraries Committee asked a number of commercial firms how much they used the National Reference Library of Science and Invention; most replied that they could not afford to send staff to London to do literature searches, but used local resources and borrowed from NLL. Thus for some commercial firms it was worth while to borrow material from NLL, but not to send staff to London to look at literature. Mr. Dammers said that the costs of maintaining a library in a firm and borrowing literature from NLL were far less than the costs involved in sending staff to a central library in London.

Mr. Mackenzie said that clearly more research was needed into library management problems and wondered what part OSTI would take in this in the future. Dr. Thompson described the structure of the recent reorganization at the Department of Education and Science and the formation of the Library and Information Systems Branch, and said that OSTI's terms of reference were the same as before. Mr. Brookes asked about the overlap of the responsibilities of OSTI and the Ministry of Technology. Dr. Thompson replied that in general terms, OSTI was responsible for research and development, whilst the Ministry of Technology's interest was in operational systems. Dr. Urquhart said that OSTI had recently established a Library Management Research Unit at Cambridge whose brief included libraries for the humanities.

Dr. Thompson said that £60m of public money was spent annually on libraries, while expenditure on research was probably only about 1% of this figure. Mr. Mackenzie said that it was easier for industrial firms to decide how much to spend on research and development, because ultimately there was a pay-off in the form of sales of new products; there was not often such a tangible pay-off for university libraries, many vice-chancellors, and even librarians, would be unwilling to sacrifice a single book for the sake of research. Nevertheless, he felt that 5% of income would not be an unreasonable price to pay for the benefits which would accrue. Dr. Thompson said that perhaps libraries would have to cut down on their services, if they wished to finance research.

University of Lancaster Library Occasional Papers

1. Some implications for library management of scattering and obsolescence, by M. K. BUCKLAND and I. WOODBURN. May 1968. Reprinted with corrections, August 1968.
Out of print. A photocopy can be supplied for 13s. 6d. plus postage on application to the Librarian.
Also available+ through the ERIC Clearing house for Library and Information Sciences as report ED 022 502.

2. An analytical approach to duplication and availability, by M. K. BUCKLAND and I. WOODBURN. June 1968.
Out of print. A photocopy can be supplied for 12s. 0d. plus postage on application to the Librarian. Also available+ through the ERIC Clearing house for Library and Information Sciences as report ED 022 516.
A revised version of this report will appear in the journal Information Storage and Retrieval in 1969.

3. Planning library services: proceeding of a research seminar held at the University of Lancaster 9-11 July 1969, sbn 901699 01 2, edited by A. G. Mackenzie and I. M. Stuart. October, 1969. 30/- post free. ≈ $3.50
Applications should be made to: Librarian, University Library, University of Lancaster, Bailrigg, Lancaster, England.

+ Applications for photocopies from the ERIC Clearing house must be made direct to:

ERIC Document Reproduction Service,
National Cash Register Company,
4936 Fairmont Avenue,
Bethesda,
Maryland,
U. S. A. 20014.

UNIVERSITY OF LANCASTER LIBRARY PUBLICATIONS

Guides

[1] Charles Baudelaire, 1821-1867: centenary exhibition. Catalogue by M.C. Pakenham. April 1967. (Out of print)

[2] Current abstracts and indexes, by Ian M. Stuart. October 1968.

[3] Russian and Soviet studies in the University Library, by G.P.M. Walker. October 1968.

4. Surrealism and Dadaism: a hand-list by L.M. Newman. September 1969.

Occasional Papers

1. Some implications for library management of scattering and obsolescence, by M.K. Buckland and I. Woodburn. May 1968. (Out of print)

2. An analytical approach to duplication and availability, by M.K. Buckland and I. Woodburn. June 1968. (Out of print)

3. Planning library services: proceedings of a research seminar held at the University of Lancaster, 9 - 11 July, 1969, ed. by A. Graham Mackenzie and Ian M. Stuart. September 1969. 30/-

Miscellaneous Publications

Report of the Librarian. 1963/4-

Guide to the Library. September 1969.*

School of Education: a guide to the services and stock of the Libraries of the University and constituent Colleges. June 1969.

Serials list. September 1969.*

Serials list. Supplement: British Government and international organisations serials. June 1969.

Ulserplan: a union list of serial publications held by libraries in Lancaster. September 1967.*

Theses for higher degrees deposited in the University Library, 1965-1968. June 1969.*

* Date of the latest revision.